

Universidade de São Paulo  
Escola Superior de Agricultura “Luiz de Queiroz”

Desenvolvimento de um método rápido de identificação, ao nível de espécie, de *Lactobacillus* e *Saccharomyces* em dornas de fermentação, por meio da técnica de MALDI-TOF MS: validação molecular e construção do banco de dados espectral

**Juliana Guimarães Fonseca**

Tese apresentada para obtenção do título de Doutora em Ciências. Área de concentração: Genética e Melhoramento de Plantas

Piracicaba  
2019

Juliana Guimarães Fonseca  
Bacharel em Biotecnologia

Desenvolvimento de um método rápido de identificação, ao nível de espécie, de *Lactobacillus* e *Saccharomyces* em dornas de fermentação, por meio da técnica de MALDI-TOF MS: validação molecular e construção do banco de dados espectral  
versão revisada de acordo com a resolução CoPGr 6018 de 2011

Orientador:

Prof. Dr. **CARLOS ALBERTO LABATE**

Tese apresentada para obtenção do título de Doutora em Ciências. Área de concentração: Genética e Melhoramento de Plantas

Piracicaba  
2019

**Dados Internacionais de Catalogação na Publicação**  
**DIVISÃO DE BIBLIOTECA – DIBD/ESALQ/USP**

Fonseca, Juliana Guimarães

Desenvolvimento de um método rápido de identificação, ao nível de espécie, de *Lactobacillus* e *Saccharomyces* em dornas de fermentação, por meio da técnica de MALDI-TOF MS: validação molecular e construção do banco de dados espectral / Juliana Guimarães Fonseca. - - versão revisada de acordo com a resolução CoPGr 6018 de 2011. - - Piracicaba, 2019.

93 p.

Tese (Doutorado) - - USP / Escola Superior de Agricultura "Luiz de Queiroz".

1. Bactérias ácido lácticas 2. Fermentação 3. Espectrometria de massas  
4. *Fingerprinting* I. Título

## DEDICATÓRIA

**Dedico este trabalho a todos que  
optaram pela pesquisa exclusivamente  
por amor e vocação.**

## AGRADECIMENTOS

Eu queria demonstrar minha completa gratidão a todos que me ajudaram de alguma forma na realização deste trabalho.

Agradeço, primeiramente, ao Prof. Labate, pela orientação, aprendizado, incentivo e carinho nesses muitos anos de convívio, por ter enxergado potencial em mim desde que eu era uma aluna de mestrado.

Agradeço à Mônica pelo constante suporte tanto profissional quanto pessoal, pelos seus dotes culinários, humildade e delicadeza. Pelas piadas, risadas, carinho, mas, principalmente, agradeço por sua generosidade em ensinar. Biologia molecular virou uma paixão também para mim.

Gostaria de agradecer especialmente a Thais, especialista do laboratório, por todo ensinamento em espectrometria de massas dado com muita paciência, por ter me incentivado quando achei que não era capaz, me apoiado e contribuído substancialmente para meu amadurecimento como pessoa e, principalmente como pesquisadora. Esse trabalho é nosso!

Eu queria agradecer a todos os integrantes da turminha Max Feffer: Felipinho, Fabrício Ana Paula, Mari, Thaís, Ilara, Andressa, Maria Letícia, Haná, Aline, Flávia, Vanessa, Lucas, Thalita e Lívia, por terem sido minha família esalquiiana, pelo convívio diário e por terem sido meu suporte nas horas que eu mais precisei.

Agradeço especialmente ao Fabrício pelas inúmeras análises de bioinformática, dúvidas no R e por ter sido um ótimo amigo todos os momentos, inclusive nas discussões sobre absolutamente tudo. À Flávia, por ter sido minha amiga-irmã desde a faculdade, passando comigo por todas as fases da minha vida, fazendo meus dias mais leves e felizes. À Mari pela sintonia, desabafos, colo, conselhos no café e cumplicidade em todas as horas. À Ana por ter sido essa pessoa maravilhosa comigo, por ficar feliz por minhas conquistas, torcer pelo meu sucesso e por me fazer sambar e rir quando as coisas não davam certo. À Maria Letícia pelas palavras de incentivo, todas as correções, dicas, cervejas e amizade.

Agradeço também ao pessoal da Bruker, principalmente ao Diego, que com toda paciência instalou os softwares e deu uma aula sobre processamento e análises dos dados do MALDI.

Agradeço especialmente a Professora Cláudia pelo trabalho prestado a pós-graduação, por ser um exemplo de dedicação, docência e profissionalismo. Cláudia você é nosso incentivo diário!

Agradeço encarecidamente ao Professor Marcos Eberlin por ter aberto as portas do Laboratório ThoMson, por ter nos recepcionado todas as vezes que precisamos e por ter permitido a realização deste trabalho. Agradeço também sua aluna, Fernanda Negrão por ter nos orientado, estando com nós em todas as aquisições e por sua imensa generosidade.

Agradeço ao Professor Basso da ESALQ por ter cedido os microrganismos para este estudo e pelas melhores aulas de bioquímica e fermentação.

Agradeço à minha família por ter me suportado nos momentos de puro estresse, por sempre verem o melhor em mim, por nunca perderem a fé quando eu pensei em desistir, obrigada.

Agradeço imensamente ao meu marido, Eduardo, por ter permanecido ao meu lado SEMPRE.

Eu gostaria de agradecer também aos professores e funcionários do departamento de genética pelo auxílio e dedicação durante esses 4 anos.

Eu agradeço a CAPES e ao CNPQ pelo suporte financeiro neste projeto.

## EPÍGRAFE

*“Creio que as aptidões necessárias a uma verdadeira vocação científica são coisas preciosas e delicadas.*

*Um tesouro raro que seria absurdo de se desperdiçar e que deve ser zelado com solicitude para que ele tenha todas as chances de florescer ”*

*Marie Curie (1867 – 1934)*

*Mãe da radiação,*

*Primeira mulher a ser laureada com um prêmio Nobel*

*E a única a ganha-lo duas vezes.*

## SUMÁRIO

RESUMO.....	8
ABSTRACT .....	9
1. INTRODUÇÃO.....	11
2. REVISÃO BIBLIOGRÁFICA .....	15
2.1. PERSPECTIVAS DA PRODUÇÃO DO ETANOL NO BRASIL.....	15
2.2. PROCESSO DE FERMENTAÇÃO PARA A PRODUÇÃO DE ETANOL A PARTIR DE CANA-DE-AÇÚCAR .....	16
2.3. CONTAMINAÇÃO BACTERIANA DO PROCESSO DE FERMENTAÇÃO: BACTÉRIAS ÁCIDO LÁCTICAS .....	18
2.4. IDENTIFICAÇÃO DE BACTÉRIAS ÁCIDO LÁCTICAS CONTAMINANTES E LEVEDURAS DO PROCESSO DE FERMENTAÇÃO .....	22
2.5. IDENTIFICAÇÃO DO PERFIL PROTEICO DOS MICRORGANISMOS CONTAMINANTES POR IONIZAÇÃO POR DESSORÇÃO A LASER ASSISTIDA POR MATRIZ (MALDI).....	24
3. JUSTIFICATIVA DO TRABALHO.....	27
4. OBJETIVOS .....	29
5. METODOLOGIA.....	31
5.1. SELEÇÃO DOS MICRORGANISMOS PRESENTES NO PROCESSO DE PRODUÇÃO DE ETANOL .....	31
5.2. IDENTIFICAÇÃO DE BACTÉRIAS ÁCIDO LÁCTICAS .....	32
5.2.1. Extração do DNA genômico total bacteriano .....	32
5.2.2. Amplificação da região 16S do rDNA.....	33
5.2.3. Purificação do produto de PCR do gene 16S do rDNA .....	33
5.2.4. Reação de sequenciamento das amostras e identificação dos microrganismos pelo gene 16S rDNA .....	34
5.2.5. Amplificação dos genes housekeeping: <i>pheS</i> e <i>groEL</i> .....	34
5.2.6. Purificação do produto de PCR do gene housekeeping <i>pheS</i> e <i>groEL</i> .....	35
5.2.7. Reação de sequenciamento das amostras e identificação dos microrganismos pelos genes housekeeping <i>pheS</i> e <i>groEL</i> .....	35
5.2.8. Análises Filogenéticas de <i>Lactobacillus</i> sp.....	36
5.3. IDENTIFICAÇÃO DE LEVEDURAS SELECIONADAS DO PROCESSO FERMENTATIVO.....	36
5.3.1. Extração do DNA genômico total de eucariotos .....	36
5.3.2. Amplificação da região parcial do gene ITS e 28S-nr LSU .....	37
5.3.3. Purificação do produto de PCR dos genes ITS e 28S-nr LSU .....	38
5.3.4. Reação de sequenciamento das amostras e identificação dos microrganismos pela região dos genes ITS e 28S-nr LSU .....	38
5.3.5. Análises Filogenéticas de <i>Saccharomyces</i> sp.....	39
5.4. PREPARO DE AMOSTRAS PARA IDENTIFICAÇÃO DOS ESPECTROS REFERENTES AOS DIFERENTES MICRORGANISMOS UTILIZANDO O INSTRUMENTO ANALÍTICO MALDI-TOF MS BRUKER MICROFLEX .....	39
5.4.1. Preparo das amostras bacterianas para aquisição por MALDI-TOF MS.....	39
5.4.2. Preparo das amostras de leveduras para aquisição por MALDI-TOF MS.....	40
5.5. AQUISIÇÃO DOS ESPECTROS E ANÁLISE DOS DADOS OBTIDOS POR MALDI-TOF MS .....	41
5.5.1. Perfil proteico espectral.....	41
5.5.2. Validação da técnica com microrganismos comerciais .....	41
5.5.3. Perfil espectral em diferentes meios de cultura .....	42
5.5.4. Regiões espectrais características de cada espécie .....	42
5.5.5. Análise estatística multivariada a partir dos dados de MALDI-TOF MS.....	42
5.5.6. Teste preliminar de identificação de leveduras a partir de uma amostra de dorna industrial .....	43
6. RESULTADOS E DISCUSSÃO.....	44
6.1. IDENTIFICAÇÃO DOS MICRORGANISMOS POR TÉCNICAS MOLECULARES .....	44
6.1.1. Crescimento em meio seletivo e seleção por microscopia óptica .....	44
6.1.2. Amplificação da região 16S rDNA de bactérias ácido lácticas .....	45
6.1.3. Amplificação dos genes housekeeping das bactérias pertencentes ao gênero <i>Lactobacillus</i> .....	46

6.1.4. Identificação molecular a nível de espécie de isolados de <i>Lactobacillus</i> por sequenciamento de regiões gênicas conservadas: 16S rDNA , <i>pheS</i> e <i>groEL</i> .....	47
6.1.5. Identificação de leveduras pelos genes <i>ITS</i> , 28S- <i>nr LSU</i> .....	54
6.1.6. Árvore Filogenética dos isolados de <i>Saccharomyces</i> .....	54
6.2. PREPARO DE AMOSTRAS PARA IDENTIFICAÇÃO DOS ESPECTROS REFERENTES AOS DIFERENTES MICRORGANISMOS UTILIZANDO O SISTEMA MALDI-TOF BRUKER MICROFLEX.....	57
6.2.1. Otimização do preparo das amostras bacterianas para aquisição por MALDI-TOF MS.....	57
6.2.2. Preparo das amostras de leveduras para aquisição por MALDI-TOF MS .....	61
6.3. AQUISIÇÃO DOS ESPECTROS DE LACTOBACILLUS E ANÁLISE DOS DADOS OBTIDOS POR MALDI-TOF MS. 62	
6.3.1. Perfil proteico espectral de isolados de <i>Lactobacillus</i> comerciais em meio MRS para validação da técnica MALDI-TOF para identificação microbiológica .....	62
6.3.2. Perfil espectral de isolados de <i>Lactobacillus</i> comerciais em diferentes meios de cultura .....	64
6.3.3. Perfil proteico espectral dos isolados de <i>Lactobacillus</i> do processo de fermentação etanólica .....	67
6.3.4. Comparação do poder discriminatório para identificação de isolados de <i>Lactobacillus</i> ao nível de espécie entre o banco de dados do Biotyper e o banco de dados interno desenvolvido .....	73
6.3.5. Análise estatística multivariada de isolados de <i>Lactobacillus</i> a partir dos dados de MALDI-TOF MS .....	75
6.4. AQUISIÇÃO DOS ESPECTROS DE SACCHAROMYCES E ANÁLISE DOS DADOS OBTIDOS POR MALDI-TOF MS .....	77
6.4.1. Perfil proteico espectral dos isolados de <i>Saccharomyces</i> .....	77
7. CONSIDERAÇÕES FINAIS .....	83
REFERÊNCIAS .....	85

## RESUMO

**Desenvolvimento de um método rápido de identificação, ao nível de espécie, de *Lactobacillus* e *Saccharomyces* em dornas de fermentação, por meio da técnica de MALDI-TOF MS: validação molecular e construção do banco de dados espectral**

O gênero *Lactobacillus* é o principal grupo de bactérias contaminantes em dornas de fermentação para produção de etanol em larga escala. A alta proliferação destes microrganismos prejudica a viabilidade de cepas de *Saccharomyces cerevisiae* selecionadas, podendo diminuir a produção de etanol nas dornas de fermentação. Os métodos mais utilizados para identificação destes microrganismos são bioquímicos e moleculares baseados na sequência do DNA, que são muito demorados, onerosos e muitas vezes remetem a resultados ambíguos. MALDI-TOF MS é uma poderosa ferramenta de identificação microbiana e foi testada neste trabalho para identificação de diferentes isolados de *Lactobacillus* e *S. cerevisiae* presentes no processo de produção de etanol. Os métodos de preparo e aplicação da amostra para aquisição espectral foram estabelecidos, tanto para bactérias quanto para leveduras. Vinte e sete isolados de *Lactobacillus* foram identificados pela região 16S rDNA e dois genes *housekeeping*, *pheS* e *groEL* e, três cepas de leveduras selecionadas do processo foram identificadas pela região ITS e 28S nr-LSU. As identificações genômicas foram contrastadas com as identificações obtidas com o perfil proteico por MALDI-TOF MS e 97% dos *Lactobacillus* tiveram a mesma classificação molecular que o gene *pheS*, eleito como o mais discriminatório, e 100% das cepas de leveduras foram classificadas como *S. cerevisiae* por ambas as técnicas. A técnica MALDI-TOF MS se mostrou altamente eficiente para discriminação intraespecífica de leveduras e interespecífica de *Lactobacillus*, não havendo apenas discriminação para isolados classificados como *L. casei* pelos genes *housekeeping*. Além disso, quando comparado o poder discriminatório da técnica MALDI-TOF MS em relação ao banco de dados espectral disponível no Biotyper e, posteriormente, ao banco de dados complementar criado neste trabalho com os microrganismos próprios do processo de produção de etanol, houve um aumento de 57 a 100% das repetições que foram identificadas ao nível de espécie com alta confiabilidade. Desta forma, os resultados deste estudo mostraram que a técnica MALDI-TOF MS pode ser utilizada como uma alternativa rápida e eficaz para identificação de *Lactobacillus* e *Saccharomyces* do processo etanólico.

Palavras-chave: Bactérias ácido lácticas; Genes *housekeeping*; Espectrometria de massas; *Fingerprinting*

## ABSTRACT

### Development of a rapid identification method at the species level of *Lactobacillus* and *Saccharomyces* in fermentation process by MALDI-TOF MS: molecular validation and spectral database construction

The genus *Lactobacillus* is the main group of contaminating bacteria in fermentation tanks for large-scale ethanol production. The high proliferation of these organisms affect the viability of selected strains of *Saccharomyces cerevisiae*, which can reduce the production of ethanol in fermentation tanks. The most used methods for identifying these microorganisms are biochemical and molecular based on DNA sequence, which are very time-consuming, costly and often revert to ambiguous results. MALDI-TOF MS is a powerful microbial identification tool and it was tested in this work to identify different isolates of *Lactobacillus* and *S. cerevisiae* present in the ethanol production process. The sample preparation and application in the MALDI plate for spectral acquisition were established for both bacteria and yeasts. Twenty-seven isolates of *Lactobacillus* were identified by the 16S rDNA region and two housekeeping genes (*pheS* and *groEL* genes) and three yeast strains selected from the process were identified by the ITS and 28S nr-LSU regions. The genomic identifications were compared with the MALDI-TOF MS protein profiles and 97% of the *Lactobacillus* had the same molecular classification as the *pheS* gene, which was chosen as the most discriminatory gene, and 100% of the yeast strains were classified as *S. cerevisiae* by both techniques. The MALDI-TOF MS technique proved highly efficient for intraspecific yeast and interspecific discrimination of *Lactobacillus*, and there was no discrimination for isolates classified as *L. casei* by housekeeping genes. Furthermore, when compared to the discriminatory power of MALDI-TOF MS technique in relation to spectral database available on Biotyper and subsequently, the complementary database created in this work with the microorganisms themselves in the ethanol production process, there was an increase from 57 to 100% of the repetitions that were identified at the species level with high reliability. Thus, the results of this study showed that the MALDI-TOF MS technique can be used as a fast and efficient alternative for the identification of *Lactobacillus* and *Saccharomyces* of the ethanolic process.

Keywords: Lactic acid bacteria; Housekeeping genes; Mass spectrometry; Fingerprinting



## 1. INTRODUÇÃO

Nas últimas décadas houve uma transformação no setor energético mundial com a necessidade da busca de fontes alternativas de combustíveis devido à finitude dos combustíveis fósseis e, também, pela independência energética necessária para o desenvolvimento dos países em ascensão, tal como o Brasil, principalmente depois da primeira crise do petróleo em 1973, levando o país a lançar o programa nacional pró-álcool (AMORIM & LEÃO, 2005). Além do contexto econômico, houve uma motivação ambiental, pois elevadas concentrações de emissões de carbono liberadas na atmosfera pela combustão dos derivados do petróleo contribuíram para o aumento do efeito estufa. Nesse contexto, o etanol ganhou espaço por se tratar de um biocombustível ambientalmente mais “amigável” e competitivo no mercado. Desta forma, o Brasil que em 1970 importava 80% do seu consumo total de petróleo, conseguiu conquistar quase que totalmente a independência energética e tornar-se líder em energia renovável. Estimativas apontam que mais de 40% da energia produzida no Brasil vem de fontes renováveis em comparação com uma média de menos de 15% do resto do mundo, sendo a cana-de-açúcar a matéria-prima que provém quase 17% da energia total do país, superando as hidroelétricas (SugarCane.org, 2018). Segundo dados da Companhia Nacional de Abastecimento (CONAB, 2018) estima-se que a safra de 2018/2019 será de 30,41 bilhões de litros de etanol, com um aumento de 11.6% em relação a safra anterior.

O processo responsável pela produção de etanol a partir da cana-de-açúcar é a fermentação, sendo a *Saccharomyces cerevisiae* a levedura mais utilizada neste processo. Este microrganismo é exaustivamente estudado e possui um vasto conhecimento acerca de seu genoma e funcionamento celular devido sua importância no processo de fermentação alcoólica (FORSTER et al., 2003). A seleção de linhagens *S. cerevisiae* eficientes e resistentes ao processo fermentativo, tais como CAT-1, Pedra -2, Santa-Adélia e Baker yeast (BASSO et al., 2008) foram de extrema importância para a alta produtividade do processo fermentativo na indústria alcooleira.

Apesar desta alta produtividade, o processo de fermentação em escala industrial não é conduzido em ambiente asséptico, ocasionando o crescimento de muitos microrganismos contaminantes que podem interferir direta ou indiretamente com o processo fermentativo. Além disso, as usinas realizam o reciclo de células de leveduras, o que ocasiona, indiretamente, o reciclo de microrganismos contaminantes, levando, a cada reciclo, a seleção de bactérias adaptadas ao ambiente fermentativo. Juntamente com a falta de qualidade e sanidade da matéria-prima utilizada, más condições de armazenamento do melão e esterilização inadequada de dornas e canalizações, propiciam o ambiente ideal para proliferação de contaminantes no processo (AMORIM, 2018).

Assim, o processo industrial de fermentação ocorre na presença de uma quantidade significativa de contaminantes bacterianos, principalmente bactérias ácido-lácticas e, quando ocorre a alta proliferação destes microrganismos, há perda na produtividade e rendimento da fermentação. A alternativa utilizada para controlar esse problema, ainda que custosa econômica e ambientalmente, é o uso de antibióticos convencionais de amplo espectro, deixando resíduos no fermento (CARVALHO; MONTEIRO, 2011).

Pela sua importância, diferentes pesquisas realizadas no Brasil tiveram por objetivo a identificação dos microrganismos contaminantes do processo fermentativo por meio de análises físico-químicas, microbiológicas e moleculares (CARVALHO; MONTEIRO, 2011; DE SOUZA E MONTEIRO, 2011; DE SOUZA et al., 2012; SOBRINHO; SILVA; CEREDA, 2011). Entretanto, uma tecnologia que vem ganhando destaque na microbiologia é a espectrometria de massas (*MS – mass spectrometry*), mais especificamente do tipo MALDI-TOF (matriz assistida por dessorção e ionização por tempo de voo, do inglês *Matrix-assisted Desorption/Ionization time-of-flight*). Apesar dessa técnica já ter sido utilizada para identificação de bactérias em 1975, foi somente décadas depois, com a descoberta de

outras matrizes que permitiram a ionização de proteínas ribossomais (range de 2- 20 kDa), seguido pelo desenvolvimento de banco de dados específicos para identificação microbiana que a técnica MALDI-TOF MS tornou-se uma alternativa efetiva na rotina clínica. A identificação é realizada a partir da obtenção de um espectro de massas característico do microrganismo por meio da ionização de proteínas intactas presentes na amostra e é comparado com perfis espectrais de outros microrganismos conhecidos presentes em banco de dados.

O primeiro sistema MALDI-TOF MS acoplado com software de dados capaz de realizar a rápida identificação microbiana em laboratórios foi desenvolvido pela Bruker Daltonics em 2008 e foi chamado de MALDI-Biotyper. Em 2008 também foram realizados diversos estudos multicêntricos mundiais validando o desempenho notável para identificação utilizando este software de dados. A plataforma MALDI-Biotyper é formada por softwares operacionais e de análises e um banco de dados espectral de microrganismos que foram submetidos a um método de extração/ preparação de amostra simples (SENG et al., 2009) e, além disso, a plataforma permite a adição de bibliotecas internas de outros microrganismos (SIGNAL et al., 2015). Esta técnica revolucionou a identificação de microrganismos por 3 motivos principais: 1) Não têm necessidade de extrair o material genético do microrganismo para a identificação, dispensando o extenso trabalho laboratorial; 2) Obtenção dos resultados dos espectros representativos quase que instantaneamente, enquanto que por meio de técnicas moleculares levaria semanas e; 3) O custo de preparo das amostras analisadas por MS é menor do que as análises e preparo de amostras feitas por sequenciamento de regiões conservadas do genoma (GARCIA et al., 2012). Porém, a identificação de microrganismos por espectrometria de massas ainda não foi utilizada no ambiente da fermentação etanólica em usinas brasileiras. No entanto, os estudos que priorizam a identificação rápida de microrganismos por espectrometria de massas, a partir de amostras complexas, já estão sendo utilizada na área médica e, mais recentemente, também para biotipagem de cepas comerciais de *S.cerevisiae* (MOOTHOO-PADAYACHIE et al., 2012), identificação de bactérias (LV et al., 2016) e leveduras (GUTIÉRREZ et al., 2017) no processo de fabricação de vinhos.

Estudos realizados para a detecção de *Enterobacteriaceae*, *Pseudomonas aeruginosa* e *Acinetobacter baumannii*, bactérias produtoras de enzimas capazes de degradar penicilina, estão sendo realizados diretamente de hemoculturas por MALDI-TOF. Essa técnica permitiu a detecção precoce da atividade dessa enzima em culturas de sangue com 119 isolados previamente caracterizados (OVIAÑO et al., 2016). Pongla e colaboradores (2016) visando a identificação de isolados bacilos gram-negativos não-fermentativos de glicose em pacientes com fibrose cística, obtiveram resultados satisfatórios com o uso de MALDI-TOF MS para a identificação de bactérias do gênero *Burkholderia* ao nível de espécie e ao nível de gênero para outros microrganismos não-fermentativos.

Lv e colaboradores (2016) utilizaram a técnica MALDI-TOF MS em conjunto com técnicas moleculares para a identificação de bactérias presentes no processo fermentativo de vinho a partir de arroz e isolaram 60 espécies diferentes de bactérias pertencentes a seis diferentes gêneros (*Bacillus*, *Staphylococcus*, *Leuconostoc*, *Pediococcus*, *Lactobacillus* e *Lactococcus*) por filogenia do gene 16S rDNA e, os espectros gerados por MALDI-TOF MS foram comparados com o banco de dados MALDI-Biotyper. Os resultados mostraram que a técnica de *fingerprinting* por MALDI-TOF MS foi complementar aos resultados obtidos pelo sequenciamento da região 16S rDNA, sendo uma técnica com maior poder discriminatório na identificação de espécies de *Bacillus*, especificamente na diferenciação entre *Bacillus subtilis*, *Bacillus amyloliquefaciens* e *Bacillus methylotrophicus*, e também em algumas espécies do gênero de *Lactobacillus*. Para a discriminação de isolados de *Lactobacillus*, especificamente, grande parte dos trabalhos não obtiveram resultados satisfatórios quando utilizaram apenas a região 16S rDNA (NASER et al., 2007; PODLESNY et al., 2011; DUSKOVÁ et al., 2012; DELAVENNE et al., 2013), mas quando utilizaram a técnica MLST (*Multilocus Sequence Typing*), sequenciando diferentes genes *housekeeping* como *pheS*, *rpoA* e *groEL*, conseguiram discriminar de maneira

eficiente a maior parte dos isolados ao nível de espécie. Delavenne et al. (2013), objetivando a identificação de potenciais isolados de *Lactobacillus* como culturas bioprotetoras no iogurte, comparou os resultados obtidos pelo sequenciamento dos genes *pheS* e *rpoA* com análises de fingerprinting utilizando MALDI-TOF MS e obteve resultados similares e satisfatórios com alta capacidade de discriminação de todos os isolados testados.

Já em relação à identificação de leveduras, em outro trabalho objetivando a identificação dos microrganismos responsáveis pela complexidade do aroma de vinhos, Gutiérrez e colaboradores (2017) utilizaram a técnica MALDI-TOF MS para identificar *Saccharomyces cerevisiae* e leveduras não-*Saccharomyces* presentes no processo fermentativo de produção de vinhos. Além disso, como a maioria dos bancos de dados de referência de espectros são quase que exclusivamente de microrganismos clínicos, Gutiérrez e colaboradores (2017) desenvolveram um banco de dados espectral complementar interno de leveduras de fermentação e, conseqüentemente, melhorou a performance de identificação.

O desafio inicial deste trabalho concentra-se na inacessibilidade de coleções apuradas de microrganismos presentes especificamente no processo de produção de etanol brasileiro, que permitam a rápida identificação utilizando a técnica MALDI-TOF MS. Na área clínica, em que grande parte dos microrganismos já foram caracterizados molecularmente e, possuindo um acervo de 7.311 espectros de referência no banco de dados Biotyper, permitiu a identificação de microrganismos diretamente do sangue (MOUSSAQUI et al., 2010), urina (FERREIRA et al., 2010) e fluido cérebro-espinhal (SEGAWA et al., 2010), sem necessidade de cultivo prévio, utilizando um kit (MALDI Sepsityper kit – Bruker, Daltonics) para lise bacteriana e lavagem da amostra antes do preparo de amostra e extração proteica simples habitual. Baseado nesses trabalhos, Barreiro (2016), avaliou a sensibilidade da técnica para detectar contaminação de diferentes espécies de bactérias causadoras de mastite (*Staphylococcus aureus*, *Streptococcus uberis*, *Streptococcus agalactiae*, *Streptococcus dysgalactiae* e *Escherichia coli*) diretamente de amostras do leite. Os resultados revelaram que a identificação dos patógenos foi possível a partir de  $10^6$  unidades formadoras de colônia (UFC)/mL para *S.aureus*,  $10^7$  UFC/mL para *E.coli* e  $10^8$  UFC/mL para as demais cepas. Apesar de ter sido possível a identificação diretamente das amostras de leite, sem necessidade de cultura prévia, houve uma perda de sensibilidade quando comparado com a identificação a partir da cultura microbiológica.

Desta forma, para realização deste trabalho, primeiramente, foi necessária a identificação de bactérias contaminantes do gênero *Lactobacillus* ao nível de espécie a partir da junção do sequenciamento da região 16S do rDNA com a análise MLST de genes *housekeeping*, servindo como base para a rápida identificação pela análise de fingerprinting utilizando MALDI-TOF. Adicionalmente, como as bactérias contaminantes vivem concomitantemente com as leveduras, em um ambiente complexo, decidiu-se testar se a técnica de identificação espectrométrica também seria eficiente para microrganismos eucariotos específicos do processo de produção de etanol brasileiro e, se era possível diferenciar os isolados de *S. cerevisiae* utilizados como microrganismos de arranque no processo de fermentação. Assim, este trabalho pioneiro propôs aplicar a técnica de “*fingerprinting*” por MALDI-TOF para a rápida identificação de isolados pertencentes ao principal gênero de bactérias contaminantes do processo de produção de etanol e diferenciar cepas de leveduras selecionadas do processo, para, posteriormente, construir um banco de dados espectral complementar que otimize e viabilize a identificação dos microrganismos presentes nas dornas de fermentação.



## 2. REVISÃO BIBLIOGRÁFICA

### 2.1. Perspectivas da produção do etanol no Brasil

No Brasil o etanol é produzido por meio da fermentação de açúcares de cana-de-açúcar por microrganismos, os quais são capazes de transformar moléculas de glicose e frutose ( $C_6H_{12}O_6$ ) em moléculas de etanol ( $C_2H_5OH$ ). A levedura *Saccharomyces cerevisiae* – organismo modelo em eucariotos – tem sido o principal microrganismo utilizado nos processos fermentativos, devido a sua excelente capacidade de crescimento e fermentação sob condições anaeróbias (FORSTER et al., 2003).

Atualmente, o Brasil tornou-se referência mundial em sustentabilidade, sendo o segundo maior produtor mundial de etanol (MAPA, 2018) estimulado pela crescente demanda de vendas de carros flex na década passada (IBETO et al.; 2011). Em produção, o Brasil fica atrás apenas dos Estados Unidos da América, o qual utiliza o milho como matéria-prima para a produção de etanol. Na safra de 2017/2018, o Brasil apresentou 9.004,7 milhões de hectares de área plantada de cana-de-açúcar, com uma produção de 651,84 mil toneladas de cana-de-açúcar, sendo que a partir disso foram produzidos 38,734 mil toneladas de açúcar e 27,254 mil metros cúbicos de etanol (ÚNICA, 2018). A região Centro-Sul foi responsável por quase 94% da produção nacional, sendo o Estado de São Paulo o maior contribuinte (56% da produção nacional total) (ÚNICA, 2018). Todo esse etanol foi produzido a partir de cana-de-açúcar e seus derivados. Estima-se que em 1000 kg de cana possua aproximadamente 160 kg de açúcar, possibilitando a produção de 79,5 kg de etanol, caso o processo atingisse 100% de rendimento (FIGUERUT, 2006; ANDRIETTA et al., 2006), pois seguindo a estequiometria da reação, a cada 100g de glicose são produzidos 64 mL de etanol, entretanto, na prática parte do carbono é destinado para o crescimento celular e produção de outros compostos como glicerol e ácido succínico, reduzindo a quantidade de etanol produzido para 61 mL /100g de glicose (MENEZES, 1980). A mensuração da perda por partição do carbono varia significamente entre os trabalhos publicados, ficando entre 10% quando o rendimento fermentativo fica em torno de 90% (BIOCONTAL, 2011 c).

Apesar da notável perda por desvio do carbono, o Brasil mantém uma produtividade considerada alta devido aos fatores climáticos favoráveis e solos férteis, a programas de desenvolvimento de novas variedades de cana-de-açúcar e do aperfeiçoamento do processo fermentativo com alto rendimento energético (LEITE et al., 2009). Apesar da base do processo de produção de etanol utilizado no Brasil ter sido desenvolvido quase 90 anos atrás por Firmino Boinot, algumas etapas do processo foram otimizadas, possibilitando um aumento no rendimento energético de 75-80% para 90% (AMORIM et al., 2011). A maior parte do etanol no Brasil é produzido por um processo de fermentação em batelada alimentada com reciclo de fermento (em torno de 80% do total da produção nacional), sendo o restante é produzido por fermentação contínua multiestágio com também reciclo de fermento (NOVACANA, 2018).

Outro fator importante para a obtenção do alto rendimento provém da seleção de linhagens *S. cerevisiae* resistentes ao processo, combinando características como: alta tolerância a pH baixo, conversão eficiente de carboidratos em etanol e tolerância à altas concentrações de etanol no meio (NEVOIGT, 2008). Algumas das linhagens comerciais conhecidas, como leveduras panificadoras (*Baker's yeast*) e algumas *Saccharomyces*: CAT-1, PEDRA-2 e SANTA-ADÉLIA são amplamente utilizadas nas destilarias (BASSO et al., 2008). Estudos realizados por Basso et al. (2008) avaliaram o desempenho de várias estirpes ao longo de 12 anos em 70 destilarias brasileiras. Os resultados mostraram que a maioria das cepas avaliadas não foram capazes de perpetuar permanentemente entre as populações de leveduras. Por outro lado, algumas foram capazes de dominar a população da fermentação a longo

prazo, entre elas as estirpes PEDRA-2 e CAT-1, as quais apresentaram as maiores capacidades de implantação, maior competitividade em relação às estirpes de contaminantes, o que representa uma média de 45 a 54% da biomassa de levedura durante a temporada de fermentação. Devido a essas características notáveis para uma fermentação industrial, CAT-1 e PEDRA-2 são ainda as cepas de arranque mais utilizadas, representando por volta de 80% da levedura seca comercializada para o etanol combustível no Brasil. Apesar da maioria das cepas utilizadas para produção de etanol serem não-floculantes em cultura pura, em condições de fermentação em larga escala, essas cepas podem apresentar características floculantes devido à co-agregação de leveduras com bactérias contaminantes, dificultando a cinética da fermentação e prejudicando a centrifugação para recuperação de células (AMORIM et al., 2011; ABREU-CAVALHEIRO E MONTEIRO, 2014).

## **2.2. Processo de fermentação para a produção de etanol a partir de cana-de-açúcar**

O processo de fermentação no Brasil (simplificado na Figura 1) ocorre em reatores de grande capacidade (0,5 a 3 milhões de litros) onde o caldo de cana ou melaço diluído, ou ainda a mistura de ambos, são utilizados como substrato, sendo depois acrescentado o fermento que é composto por leveduras em alta densidade celular (10-15% v/v). Desta forma, o processo de fermentação é curto, cerca de 6-12 horas à 32-35°C, sendo o vinho depois centrifugado para recuperação das leveduras, as quais serão utilizadas novamente em um novo ciclo de fermentação. As leveduras são reutilizadas durante toda safra da cana-de-açúcar, sendo que no final de cada ciclo de fermentação (8 a 10h) elas são tratadas com ácido sulfúrico (pH 2.0–2.5) para diminuição da contaminação, principalmente bacteriana (AMORIM et al, 2011). Outras alternativas também podem ser utilizadas para controlar a contaminação, como a pasteurização do melaço e caldo da cana e o uso de antibióticos. A pasteurização pode reduzir a contaminação a níveis mais baixos (cerca de  $10^3$  células/mL), entretanto, devido ao alto custo, torna-se inviável para o processo. Já o uso de antibióticos é recorrente nas usinas na fase de pré-fermentação, apesar do custo elevado e dos possíveis impactos ambientais decorrentes do uso exacerbado (SILVA-FILHO et al., 2005), pois um dos subprodutos e/ou resíduos mais produzidos no processo é a vinhaça ( a cada 1L de etanol produz cerca de 15 L de vinhaça), a qual contém macro e micronutrientes extraídos da cana, que são utilizados na prática de fertirrigação, retornando ao solo (NOLASCO JUNIOR, 2010). Um agravante desta prática é que como se produz uma quantidade imensa de vinhaça, os critérios para deposição no solo são econômicos e não nutricionais, depositando no solo mais nutrientes do que o que ele é capaz de reciclar, já que o custo de transportar a vinhaça para toda a área de cana colhida é muito alto para a usina. Desta forma, muitas vezes são aplicadas sobredosagens de vinhaças no solo em áreas mais próximas ao processo, juntamente com resíduos de antibióticos, proporcionando episódios de salinização e contaminação dos lençóis freáticos (NOLASCO JUNIOR, 2010).

Estudo visando melhoria do desempenho energético de uma Usina no Mato Grosso ranqueou os principais insumos com custo impactante durante os quatro primeiros meses da safra de 2011. Oito insumos foram selecionados, dentre eles: ácido sulfúrico (8 g/ m<sup>3</sup> de álcool produzido), dispersante (0,24 g /m<sup>3</sup> de álcool produzido) e antibiótico (10,6 g/ m<sup>3</sup> de álcool produzido), todos utilizados para controle da contaminação bacteriana, os quais totalizaram mais de 65% dos custos de todos insumos industriais, representando um gasto de quase 1 milhão e meio de reais (MEDEIROS, ALMEIDA e THOMAZINI, 2012).

Já outro estudo em uma usina do interior de Goiás que opera em batelada alimentada e utiliza ácido sulfúrico e dióxido de cloro ao invés de antibióticos para controle da contaminação, calculou a influência da contaminação bacteriana na eficiência fermentativa e o custo desses insumos para os vencimentos da usina durante

as safras de 2014, 2015 e 2016. Notou-se que com o aumento do consumo dos insumos, principalmente dióxido de cloro, a contaminação não atingiu o ponto crítico de  $1,0 \times 10^7$  bastonetes/mL, diminuindo a perda na fermentação, aumentando o rendimento fermentativo e, conseqüentemente a eficiência industrial tendeu a aumentar. A eficiência industrial foi calculada por meio da porcentagem de Açúcares Redutores Totais (ART) presentes na cana-de-açúcar moída que foi de fato convertida em etanol no final do processo, levando em consideração eventuais perdas durante a extração, tratamento do caldo, fermentação e destilação. Já o rendimento fermentativo foi calculado por balanço de massa dorna a dorna a partir da determinação da quantidade de ART presente no mosto que entrou na fermentação por medidores de vazão e amostrador contínuo e a quantidade de etanol presente no vinho. Os resultados do estudo mostraram que a safra de 2015 e 2016 tiveram uma diminuição da perda de fermentação de 0,36 e 1,42%, respectivamente, em relação ao ano de 2014, o que impactou um aumento no rendimento fermentativo de 0,34% em 2015 e de 1,40% em 2016. Financeiramente, esse ganho representou um aumento na produção de etanol de 1.480 m<sup>3</sup> em 2014 e 3.807 m<sup>3</sup> em 2015, totalizando um adicional de R\$ 2.652.160,00 milhões e R\$ 6.822.144,00 milhões em 2014 e 2015, respectivamente, quando houve um uso maior de insumos para controle da contaminação bacteriana. Em 2014, gastou-se com ácido sulfúrico e dióxido de cloro R\$ 902.788,44 mil e em 2015, R\$ 126.475,95 mil, lucrando R\$ 1.749.371,56 milhões e R\$ 6.695.668,05 milhões, respectivamente, comprovando, assim, a rentabilidade do controle da contaminação na fermentação (OLIVEIRA e SILVA, 2016).

Apesar da existência de medidas visando o controle da contaminação bacteriana, a persistência desses microrganismos é notável. Segundo Skinner-Nemec et al. (2007), as bactérias contaminantes conseguem formar biofilmes em condições laboratoriais e, possivelmente mantém essa capacidade em processos em larga escala. A formação de biofilmes dificulta a eficiência dos antibióticos e limpeza das dornas. Linhagens de *Lactobacillus* têm se mostrado menos suscetível a ação de antibióticos quando há formação de biofilmes (STEWART et al., 2004). Assim, temos que independente de qual tipo de estratégia adotada e dos processos – batelada-alimentada ou contínuo – vários aspectos podem afetar direta e indiretamente o rendimento do processo fermentativo. Dentre esses aspectos, destacam-se: pH, temperatura, concentração do inóculo, concentração de produtos secundários no meio como o etanol (BASSO et al., 2008), floculação (TIUKOVA, EBERHARD e PASSOTH, 2013) e contaminação bacteriana (BASSO et al 2014). Basicamente, como cada usina possui peculiaridades nas etapas do processo de fabricação do etanol, tem-se uma seleção individualizada da biota microbiana. Assim, diferentes grupos de leveduras e bactérias são constantemente selecionadas de acordo com as condições do processo utilizado, gerando uma microbiota específica em cada fermentação, por este motivo é tão difícil prever e dimensionar as perdas de produção de etanol.

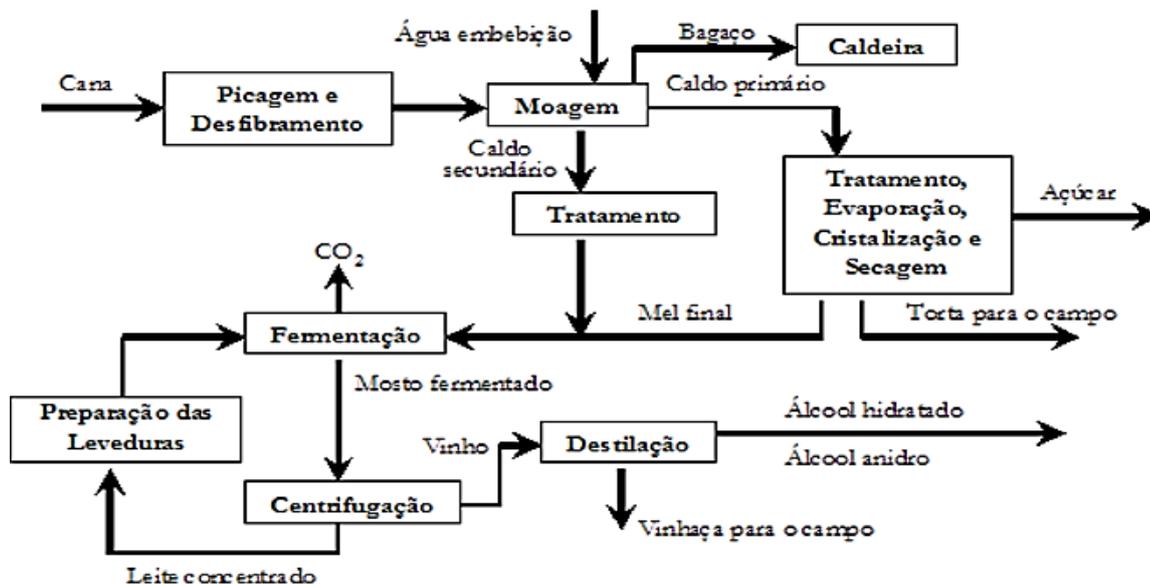


Figura 1. Ilustração simplificada das diferentes etapas dos processos que ocorrem nas usinas, mostrando desde a colheita da cana até a produção de bioetanol (álcool anidro e hidratado) e açúcar, assim como todos os subprodutos do processo (MAGALHÃES, 2007).

Nos principais tipos de processos mencionados, contínuo e batelada alimentada com reciclo de células, tem-se uma vazão constante de substrato, permitindo que sempre haja uma quantidade determinada de açúcar para ser fermentado. Com isso, até o final da fermentação, haverá carbono disponível para ser transformado em etanol pelas leveduras, mas também haverá carbono para ser desviado pelas bactérias e transformado em lactato e acetato. Notamos então que o ambiente fermentativo permite a seleção indiscriminada de microrganismos capazes de fermentar hexoses em meio anaeróbico, e não apenas a seleção de leveduras selecionadas, prevalecendo um processo de evolução conjunta da microbiota. Além disso, o ambiente da dorna modifica-se ao longo do processo fermentativo, variando de um ambiente com alto teor de substrato até um ambiente com alto teor de etanol, ácido láctico e acético. Isto faz com que a microbiota seja dinâmica e interativa. Estudos da comunidade microbiológica do processo em diferentes usinas desenvolvido por Basso et al. (2008) mostrou que na maioria das vezes, leveduras dominantes no início do processo desapareciam ao longo das safras, predominando ao final do processo, leveduras selvagens, assim como as bactérias contaminantes.

Apesar de certa individualidade de cada processo de produção de etanol e uso de diferentes matérias primas, como milho e cana-de-açúcar, todos trabalhos evidenciaram que os microrganismos contaminantes prevalentes sempre são bactérias ácido lácticas, particularmente do gênero *Lactobacillus*, representando mais de 40% dos isolados (SKINNER; LEATHERS, 2004; BONATELLI et al., 2017).

### 2.3. Contaminação bacteriana do processo de fermentação: bactérias ácido lácticas

A contaminação bacteriana começa a causar prejuízos quando atinge níveis superiores a  $1,0 \times 10^7$  células/mL de meio fermentativo (AMORIN et al., 1981). Estudos feitos por Mendes e colaboradores (2007) em cana-de-açúcar identificaram a presença de isolados dos gêneros *Pantoea*, *Pseudomonas* e *Microbacterium* por BOX PCR. E posteriormente, Magnani e colaboradores (2010) mostraram que parte do levantamento da diversidade microbiana em folhas e colmos de cana-de-açúcar foram compostos por *Enterobacteriaceae*, *Bacilli*, *Pseudomonadaceae* e *Curtobacterium*

pela amplificação do gene 16S. Além da cana in natura, o melaço utilizado como matéria-prima na fermentação pode ser fonte de contaminação do processo. Pillai e colaboradores (2011) estudou a diversidade microbiana presente no melaço e identificou bactérias dos gêneros *Pseudomonas*, *Bacillus*, *Escherichia*, *Satphylococcus* e *Micrococcus*. Outros estudos anteriores isolaram bactérias contaminantes do processo fermentativo (BISCHOFF et al., 2008), tais como: *Pediococcus*, *Enterococcus*, *Acetobacter*, *Gluconobacter* e *Clostridium* e *Lactobacillus*.

A presença de contaminantes pode interferir no processo fermentativo, seja pela disputa de nutrientes ou pela liberação de compostos secundários no meio (BASSO et al., 2014). Assim, os carbonos provenientes das moléculas de glicose e frutose disponíveis no meio fermentativo, que potencialmente poderiam ser transformados em moléculas de etanol pela levedura *Saccharomyces cerevisiae*, é capturado por bactérias contaminantes, sofrendo um desvio na rota metabólica. Este desvio, além de reduzir diretamente a produção de etanol, acarreta na produção de outros metabólitos como ácido láctico e acético, os quais prejudicam a viabilidade e performance das leveduras fermentativas, causando episódios de floculação e prejudicando a produtividade etanólica indiretamente. Por isso, a contaminação bacteriana é considerada o maior entrave do processo de produção de etanol no Brasil (AMORIM e OLIVEIRA, 1982; BAYROCK e INGLEDEW, 2004; BASSO et al., 2013).

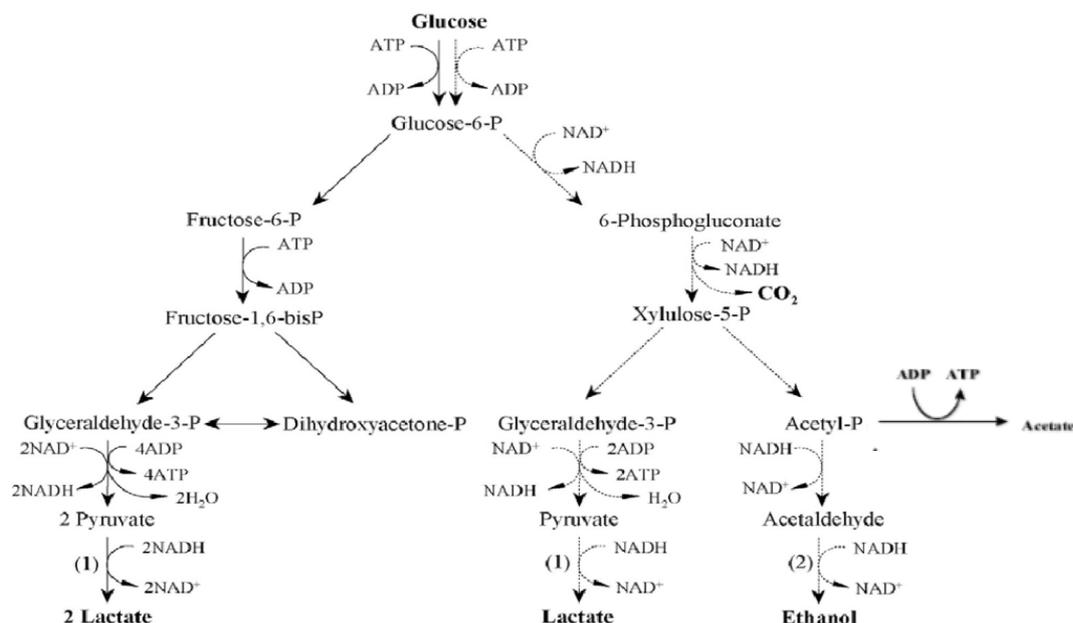
A maioria dos contaminantes bacterianos presentes no processo fermentativo são bactérias ácido lácticas (GALLO, 1990; LEUCENA et al., 2010), principalmente devido ao fato desses microrganismos terem alta tolerância a pH ácido (pH 2,0 – 3,0) e a alta concentração de etanol (8 – 11% v/v) em comparação aos outros microrganismos (KANDLER 1983; SKINNER e LEATHERS 2004; LEUCENA et al., 2010).

Estudos anteriores evidenciaram que o principal gênero prevalecente nas dornas em processo industrial são *Lactobacillus*, (GALLO, 1990; LEUCENA et al., 2010), sendo *Lactobacillus fermentum* e *Lactobacillus brevis* reportados como os contaminantes mais comuns no processo de produção de etanol (WATANABE, Nakamura e Shima, 2008). Os *Lactobacillus* possuem uma capa proteica, constituída por floculinas, que por ter uma natureza gelatinosa permitem a fixação mecânica das células de leveduras, levando a episódios de floculação (GALLO e CANHOS, 1991). Floculação é a formação de agrupamentos compostos por bactérias e leveduras, ou ainda leveduras com leveduras, o que leva a uma diminuição da velocidade cinética da fermentação, pois os agrupamentos ficam mais densos, acarretando perda de células de leveduras no fundo das dornas e dificultando a centrifugação para o reciclo de células (AMORIM; OLIVEIRA, 1982; OLIVA-NETO; YOKOYA, 1997). Segundo alguns estudos (YOKOYA e OLIVA-NETO, 1991; TIUKOVA, EBERTHARD E PASSOTH, 2014), *L. fermentum*, *L. plantarum* e *L. vini* são os principais responsáveis pelos episódios de co-agregação celular, devido á presença de uma adesina específica de manose (*Manose Specific Adhesin*- Mas) que facilita a interação célula-a-célula com leveduras.

Os *Lactobacillus* isolados no processo são classificados em dois sub-grupos metabólicos: homo e heterofermentativos (COSTA et al., 2008), de acordo com a via utilizada para metabolizar os açúcares (KANDLER, 1983) (Figura 2). As bactérias homofermentativas assimilam a glicose por meio do processo denominado glicólise, onde 1 mol de hexose resulta em 2 moles de ácido láctico e 2 moles de ATP. Entretanto, outra via é ativada nas bactérias heterofermentativas (KENDLER e WEISS, 1986), na qual 1 mol de hexose é convertido em quantidades equimolares de ácido láctico, etanol ou acetato e dióxido de carbono, produzindo 1 mol de ATP/ mol de hexose fermentada ( COGAN e JORDAN, 1994). Entretanto, quando o acetil fosfato é convertido em acetato ao invés de etanol, há acréscimo na produção de ATP, porém há necessidade de um receptor de elétrons alternativo.

Se notarmos na Figura 2, para a glicose ser transformada em xilose-5-fosfato, precisou-se que dois NADH fossem oxidados a  $NAD^+ + H^+$ . Quando há produção de lactato e etanol por microrganismos homofermentativos, os dois NADH são reduzidos a  $NAD^+$ , estabelecendo o balanço de elétrons. Entretanto, quando há formação de

lactato e acetato ao invés de etanol por microrganismos heterofermentativos, sobra um NADH a ser reduzido. Sob condições aeróbicas, a molécula de oxigênio serviria como um receptor final de elétrons (CONDON, 1987), processo denominado de respiração celular. Porém, em condições anaeróbicas, onde há uma limitação na quantidade de oxigênio, a frutose pode ser reduzida em manitol pelas bactérias heterofermentativas (VON WEYMARN et al., 2002). Isto explica o motivo pelo qual a frutose é considerada como um importante co-substrato para lactobacilos heterofermentativos e o manitol é considerado um indicador muito sensível da presença de microrganismos contaminantes no processo industrial (STEINMETZ et al., 1998; EGGLESTON et al., 2007). Outro fator interessante, é que estudos feito por Basso e colaboradores (2006) mostraram que o manitol é o segundo metabólito bacteriano mais abundante, perdendo apenas para o lactato, representando grande parte do desvio do açúcar.



**Figura 2.** Vias metabólicas de fermentação da glicose por bactérias ácido lácticas homo (esquerda) e heterofermentativas (direita), em que: P, fosfato; ADP, adenosina 5'-difosfato; ATP, adenosina 5'-trifosfato; NAD<sup>+</sup>, nicotinamida adenina dinucleotídeo (forma oxidada); NADH, nicotinamida adenina dinucleotídeo (forma reduzida); (1), lactato desidrogenase; (2), álcool desidrogenase. Figura adaptada de Wee et al., 2006.

Como a sacarose é o açúcar predominante do caldo extraído da cana, ambos os substratos, glicose e frutose apresentam proporção de 1:1. Assim, o manitol pode prever perdas de sacarose devido à síntese de dextrano bacteriano, resultando em problemas como viscosidade, evaporação, cristalização (EGGLESTON et al., 2004). Outro indicador de contaminação confiável do processo fermentativo é o ácido láctico (ALVES 1994; NARENDRANATH et al., 1997), pois é produzido por ambas bactérias, homo e heterofermentativas.

Mais recentemente, a predominância do gênero *Lactobacillus* também foi evidenciado nos estudos feitos por Lucena e colaboradores (2010) por meio da análise de amplificação do DNA ribossomal, sendo que o maior número espécies de BALs foi evidenciado nos primeiros 30 dias do início da fermentação e após 60 dias houve uma predominância das espécies *L. fermentum* e *L. vini*, correspondendo aos isolados mais abundantes nesse estudo e mais adaptados ao processo fermentativo. Notou-se também uma tendência na redução do número de espécies BALs no final do processo, sugerindo a ocorrência de resistência a antibióticos e/ou a ocorrência de infecções endêmicas. Desta forma, os autores sugeriram que as condições severas do processo (antibióticos, temperatura elevada, pH

baixo, e alta concentração de etanol) eventualmente exercem uma grande pressão seletiva sobre a microbiota da fermentação, levando a uma posterior seleção de alguns tipos BALs resistentes, predominantemente de algumas espécies de lactobacilos .

Além das BALs, bactérias produtoras de ácido acético (*Acetobacter orientalis* e *Acetobacter pasteurianus*) bactérias produtoras de amilase (*Bacillus subtilis*, *Bacillus circulans*, *Bacillus amyloliquefaciens* e *Bacillus sporothermodurans*) e bacterias patogênicas em plantas (*Burkholderia ubonensis*, *Ralstonia solanacearum* e *Pelomonas puraquae*) também foram encontradas (LUCENA et al., 2010). A presença dos microrganismos contaminantes pode afetar de maneira distinta a fermentação da *S. cerevisiae*, por exemplo, a presença de *L. fermentum* pode ocasionar floculação das células de leveduras (YOKOTA e OLIVIA NETO, 1991) em sinergia com a presença de cálcio, o que leva a perda de viabilidade das leveduras (LUDWIG, OLIVA-NETO e ANGELIS, 2001). Além disso, os autores notaram que o rendimento, o consumo de açúcar e produção de etanol foram severamente afetados quando a concentração de ácido láctico foi superior a 4,8 g/L. Já a presença de *Lactobacillus paracasei* afetou a viabilidade das células de levedura quando a concentração de ácido láctico no processo excedeu 8 g/L (NARENDRANATH, THOMAS e INGLEDEW, 1994). A viabilidade da levedura *S. cerevisiae* também foi prejudicada pela inoculação das bacterias *L. fermentum* e *L. delbrueckii* na fermentação em batelada em trigo (CHIN e INGLEDEW, 1994), provando que a presença de contaminantes bacterianos podem interferir negativamente na população de leveduras comerciais na fermentação (NOBRE, HORRI e ALCARDE, 2007).

Estudos realizados por Basso e colaboradores (2014) relataram os efeitos de BALs homo e heterofermentativas sobre o desempenho fermentativo de leveduras comerciais, como a *S. cerevisiae* - CAT1. Quando simulado as condições industriais em dornas de fermentação, a linhagem heterofermentativa de *L. fermentum* foi mais prejudicial quando comparado com a linhagem homofermentativa *L. plantarum* em relação à produção de etanol e competição com leveduras no reciclo celular.

Ainda há alguns estudos que evidenciaram a correlação entre a presença de BALs e o crescimento da levedura *Dekkera bruxcellensis* (PASSOTH, BLOMQUIST e SCHNIRER, 2007; LUCENA et al., 2010) indicando uma ecologia microbiana complexa no processo de produção de bioetanol. Entre eles, o trabalho de Souza e colaboradores (2012) avaliou os efeitos desses contaminantes sobre a viabilidade de *S. cerevisiae* e rendimento de etanol e os autores notaram que na adição de *L. vini* à culturas puras de *S. cerevisiae* não houve alteração significativa na fermentação, entretanto quando *L. vini* foi adicionado à cultura de *D. bruxcellensis* houve estímulo na atividade fermentativa da levedura, proporcionando um aumento no rendimento de etanol.

Assim, devido a coexistência entre leveduras e bactérias em meio altamente seletivo, é difícil compreender se a relação entre estes microrganismos trata-se de parasitismo ou mutualismo. Estudos apontam que ambos os microrganismos se beneficiam dessa interação (GOMES, 2009), ou seja, interação mutualística, pois mesmo em experimentos em laboratórios utilizando apenas leveduras, na ausência ou quase não presença de contaminantes, não há um aumento substancial no rendimento da produção de etanol. Além disso, quando ambos microrganismos estão crescendo concomitantemente em meio seletivo e controlado, no qual o crescimento da levedura é restritivo por condições limitantes de vitaminas, algumas substâncias essenciais para o crescimento de espécies de *Lactobacillus*, que estão ausentes no meio, tais como: lisina, triptofano, alanina, guanina, ácido nicotínico, ácido aspártico são produzidos e liberados no meio pelas células das leveduras (CHALINOR; ROSE, 1954). Desta forma, a levedura aparenta ser o microrganismo ativo, responsável por fornecer nutrientes essenciais para a proliferação de bactérias.

Entretanto, as relações antagônicas entre essas duas classes de microrganismos é determinada em grande parte pela liberação de diferentes ácidos orgânicos. Enquanto a liberação de ácido láctico e acético pelas bactérias

afetam negativamente a viabilidade das leveduras (OLIVA-NETO; YOKOYA, 1994; MAIORELLA et al., 1983; GOMES, 2010), a produção de ácido succínico pelas leveduras reprime o crescimento bacteriano (BASSO et al., 1997). Experimentos realizados com fermentação em batelada de 6 ciclos com mosto provindo de milho na presença de *S. cerevisiae* e *L. fermentum* (THOMAS et al., 2001), evidenciou que inicialmente, a presença do contaminante descreveu em 22% a produção de etanol, havendo um desvio dos carbonos para a produção de glicerol pela levedura e de ácido láctico e acético pela bactéria, diminuindo a viabilidade celular da levedura em 55%. Posteriormente, com o aumento da produção de ácido succínico pela levedura ocasionou uma diminuição da produção do ácido láctico e acético pela bactéria, diminuindo a população de microrganismos contaminantes. Ou seja, a levedura foi capaz de exercer atividade antibiose suficiente para controlar a proliferação de bactérias contaminantes. Narendranath e colaboradores (1997) testaram a influência de diferentes espécies de *Lactobacillus* (*L. paracasei*, *L. plantarum*, *L. rhamnosus* e *L. fermentum*) no desempenho fermentativo de *S. cerevisiae*, quando eram inoculadas nas concentrações  $10^5$ ,  $10^6$ ,  $10^7$ ,  $10^8$  e  $10^9$  (UFC/ mL). Eles observaram que espécies homofermentativas produziram mais ácido láctico

(1 a 1.5 % (p/v)) do que a espécie *L. fermentum*, heterofermentativa (0.5 % (p/v)). Entretanto quando as espécies *L. fermentum*, *L. plantarum* e *L. rhamnosus* foram inoculadas na concentração de  $10^6$  (UFC/ mL) houve uma diminuição de 2 % na concentração final de etanol em relação à situação controle, enquanto que para a espécie *L. paracasei*, na concentração de  $10^5$  já foi o suficiente para ter essa mesma redução. Acredita-se que esse efeito mais dramático dos isolados *L. paracasei*, seja devido a sua maior tolerância a presença de etanol, adaptação mais rápida, com menor período de defasagem e ao seu rápido crescimento. No entanto, quando a concentração celular foi de  $10^9$  de UFC/ mL, independente da espécie, houve uma redução de 3.8 a 7.6% na concentração final de etanol. Entretanto, de acordo com Nobre (2005) a presença isolada dos metabólitos de várias bactérias testadas não foram suficientes para exercer papel negativo sobre a viabilidade de leveduras, mostrando que os fatores mais agravantes para a diminuição da população de leveduras eram concomitantemente: presença de um pH ácido e disputa por nutrientes devido a presença de bactérias contaminantes.

#### **2.4. Identificação de bactérias ácido lácticas contaminantes e leveduras do processo de fermentação**

O grupo denominado bactérias do ácido lácticas foi descrito primeiramente por Beijerinck (1901). O gênero foi dividido e caracterizado inicialmente de acordo com suas características fenotípicas, temperatura de crescimento e a forma de fermentação por Orla-Jensen (1919), e posteriormente, o mesmo autor utilizou das características morfológicas, metabólicas e fisiológicas como critérios para estabelecer a sua diferenciação taxonômica. BALs são caracterizadas como microrganismos gram-positivos, catalase negativos, citocromos ausentes, não esporogênicos, aerotolerantes, exigentes quanto aos fatores nutricionais, estritamente fermentativos com produção de ácido láctico a partir da fermentação e hexoses, capazes de crescer em meio salino e tolerantes a meio ácido (YOKOYA et al., 1997; KANDLER, 1983; CARR et al., 2002). Além disso, estas bactérias são essencialmente mesófilas (desenvolvem-se em temperaturas moderadas), raras exceções em que elas são termófilas, podendo sobreviver em ambientes com ampla variação de temperatura (5 a 45°C).

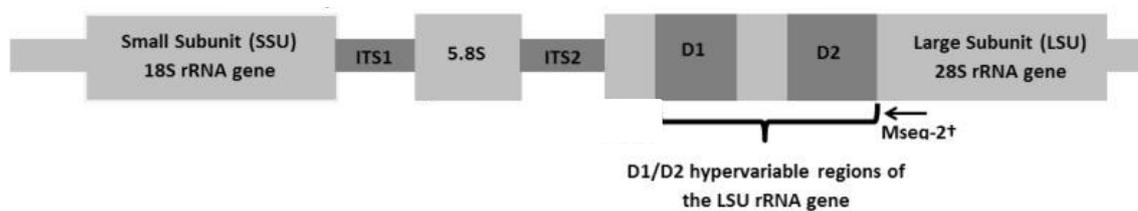
Estudos taxonômicos e de variabilidade genética em microrganismos usam frequentemente o gene codificador do RNA ribossômico, por este apresentar múltiplas cópias, ser altamente conservado e também estar

presente em todos os organismos (GUARRO et al., 1999). O sequenciamento desse gene permite a identificação de microrganismos em nível de gênero e em alguns casos ao nível de espécie (CHENEY et al., 2000). A diversidade das bactérias cultivadas pode ser acessada por meio do sequenciamento do gene ribossomal 16S, a partir da realização de PCR com primers universais bacterianos (TRINGE; HUGENHOLTZ, 2008; PACE, 1997). No entanto, apenas a região 16S do rDNA não permite a identificação precisa ao nível de espécie das bactérias devido ao alto grau de polimorfismo e de múltiplas cópias desse gene, possuindo um baixo poder discriminatório entre espécies (VASQUEZ et al., 2005). Uma alternativa para superar essa limitação é o uso da técnica MLST baseada na similaridade de conjunto de genes bem conservados – por exemplo, genes *housekeeping*- para a definição e classificação taxonômica das espécies (NASER, 2005). Dessa forma, a análise multigenica permite a distinção entre espécies que são altamente relacionadas, o que não seria possível somente com base no sequenciamento do gene ribossomal 16S.

Para o uso da técnica MLST é necessário que os genes candidatos selecionados possam representar genomas inteiros, serem moderadamente conservados, essenciais para o metabolismo bacteriano e exclusivo dentro do genoma (ZEIGLER, 2003). Assim, é necessário escolher genes *housekeeping* apropriados, pois não é possível utilizar o mesmo conjunto de genes em todos os organismos (OLIVEIRA, 2010).

Para a distinção de espécies de BALs, tais como do gênero *Lactobacillus* (NASER et al., 2007), *Carnobacterium* (SNAUWAERT et al., 2013) e *Enterococcus* (NASER et al., 2005) os genes *pheS* (subunidade alfa da fenilalanina) e *rpoA* (subunidade alfa da RNA polimerase) têm provado serem eficientes marcadores genômicos (NASER et al., 2007; De VUYST e VANCANNEYT 2006, De BRUYNE et al., 2010, 2008a, b, 2007). A análise concomitante desses dois genes proporcionou uma identificação rápida e confiável de diferentes espécies de BALs, apresentando um poder discriminatório superior ao da análise 16S rDNA (NASER et al., 2005). Outro gene também tem se mostrado um marcador confiável para a identificação e qualquer espécie de lactobacilos, o gene *groEL* (hsp60 ou 60-kDa proteína de choque térmico) (BLAIOTTA et al., 2007). O gene *groEL* codifica a proteína de choque térmico 60-kDa que está envolvida desde a síntese da cadeia polipeptídica até a formação de domínios complexos, importante para manutenção da atividade celular. Estudos utilizando este gene (DELLAGLIO et al. 2011 e BLAIOTTA et al. (2007) para identificação de bactérias lácticas, identificou cepas de *Lactobacillus* fenotipicamente e genotipicamente relacionadas.

Já para a identificação de levedura, a sequência do gene do RNA ribossômico nuclear já tem sido utilizada como alternativa ao marcador mitocondrial, especificamente ao fragmento do gene da citocromo oxidase I, há mais de 20 anos (BEGEROW et al., 2010). O gene do rRNA é composto pelos genes 18S, 5.8S e 28S transcritos em uma unidade da RNA polimerase I. Por se tratar de microrganismos eucariotos, ocorre processamento pós-transcricionais, removendo dois espaçadores internos, dividindo o gene. A região ITS é formada por esses dois espaçadores, incluindo o gene 5.8S e, foi descrita por Schoch e colaboradores (2012) como a região mais discriminatória para espécies do reino *Fungi*, juntamente com a subunidade ribossomal maior (LSU), 28S. Para leveduras, as regiões divergentes hipervariáveis na subunidade ribossômica maior, denominadas D1/D2 era adotada para caracterizar diferentes espécies, antes até do conceito de DNA barcoding ser amplamente divulgado (FELL et al., 2000; KURTZMAN e ROBNETT, 1998; SCORZETTI, FONSECA e STATZELL-TALLMAN, 2002). Ambas as regiões D1/D2 e ITS contêm regiões com sequencias conservadas e variáveis, que permitem o uso de primers universais que amplificam sequências com variabilidade suficiente para distinguir diversas espécies de fungos. A exemplificação da estrutura do gene do RNA ribossômico, com todas suas subunidades e regiões podem ser visualizadas na Figura 3.



**Figura 3.** adaptada de Arbefeville, Harris e Ferrieri (2017): Diagrama representativo do complexo gênico ribossômico composta pela região codificadora do gene 18S rRNA, pela região interna transcrita 1 (ITS1), uma região codificadora do gene 5.8S rRNA, outra região ITS denominada ITS2 e a seqüência codificadora do gene 28S rRNA. Dentro da subunidade maior do ribossomo, encontra-se a região hipervariável denominada D1/D2.

## 2.5. Identificação do perfil proteico dos microrganismos contaminantes por Ionização por dessorção a laser assistida por matriz (MALDI).

A técnica de ionização/dessorção a laser assistida por matriz, (em inglês: matrix-assisted laser desorption/ionization - MALDI) foi descrita por Karas e colaboradores em 1985. Nesta técnica, a amostra contendo a espécie de interesse na presença de uma matriz (composto necessário para a ionização da amostra) é solubilizada em solvente orgânico e/ou água e uma pequena quantidade de cada solução (normalmente 1µL) é depositada sobre uma placa. Com a evaporação do solvente, a matriz cristaliza e as moléculas da amostra se encontram inclusas nestes cristais.

A matriz utilizada é geralmente composta por ácidos orgânicos aromáticos, os quais absorvem no mesmo comprimento de onda do laser que incide sobre essa mistura. Desta forma, a energia do laser é absorvida pela matriz, o que promove a sua dessorção. As moléculas do analito, que estavam inclusas na matriz, são arrastadas para a fase gasosa e a formação dos íons ocorre por meio da transferência de prótons das moléculas da matriz para o composto analisado que é detectado na forma protonada  $[M+H]^+$ . Os íons de proteínas são eletrostaticamente acelerados até o tubo de vôo a uma velocidade diretamente proporcional à sua massa. Desta forma, íons de diferentes massas são detectados em diferentes períodos de tempo. A velocidade é calculada precisamente na faixa de nanossegundos pela diferença entre o tempo da aceleração pulsada e a detecção do sinal, o qual é posteriormente convertida em massa molecular precisa. Como o espectro de massa correspondente é específico da espécie para grande parte dos microrganismos, é utilizado como uma ferramenta de “impressão digital molecular”, podendo ser utilizada para identificação de microrganismos desconhecidos pela comparação dos espectros com banco de dados espectrais de referência (MANUAL BIOTYPER)

MALDI-TOF é uma técnica rápida e sensível para a caracterização de microrganismos (FENSELAU, 1997; FENSELAU e DEMIREV, 2001), tais como diferenciar bactérias gram-positivas de gram-negativas, cianobactérias, fungos e protozoários em diferentes níveis taxonômicos (DICKINSON et al., 2004; DIECKMANN et al., 2005; FASTNER et al., 2001; ISHIDA et al., 2003; LI et al., 2000; VAN BAAR, 2000). Corroborando com essa informação, trabalhos utilizando MALDI-TOF têm mostrado ser uma ferramenta eficaz para diferenciar bactérias em nível de gênero, espécie e, em alguns casos, até de linhagem (SANDRIN, GOLDSTEIN e SCHUMAKER, 2013; BENAGLI et al. 2012).

A técnica analítica MALDI-TOF ioniza as moléculas biológicas (tipicamente proteínas) a partir de superfícies celulares, membranas intracelulares e ribossomos de células inteiras ou a partir de extratos proteicos (SANDRIN, GOLDSTEIN e SCHUMAKER, 2013; SAUER e KLIEM, 2010; ZHANG et al., 2015). Os picos resultantes formam padrões característicos ou "impressões digitais" dos microrganismos analisados (ZHANG et al., 2015). Duas das estratégias mais utilizadas para a identificação são baseadas em bibliotecas ou em bioinformática (SANDRIN, GOLDSTEIN e SCHUMAKER, 2013). Abordagens baseadas em bibliotecas envolvem a comparação de espectros de massas desconhecidos com espectros de massas previamente coletados de microrganismos na base de dados/bibliotecas de identificação. Softwares comercialmente disponíveis, tais como BioNumerics, Biotyper e pacotes de software SARAMIS, bem como ferramentas de software de acesso aberto, como o SpectraBank (<http://www.spectrabank.org>), estão disponíveis para facilitar a comparação e quantificação de semelhança entre os espectros de massas desconhecidos com os de referência de bactérias, por exemplo. Abordagens baseadas em bioinformática geralmente contam com genomas sequenciados de microrganismos para identificação. Em contraste com a abordagem baseada em bibliotecas que não envolvem a identificação da natureza biológica de picos específicos, abordagens baseadas em bioinformática envolvem identificação das proteínas.

A biblioteca amplamente utilizada é a do MALDI-Biotyper, na qual microrganismos desconhecidos são identificados a partir da comparação da sua impressão digital com a impressão de milhares de microrganismos pertencentes ao banco. Essa comparação é realizada a partir de diferentes parâmetros. Primeiramente, o espectro do microrganismo desconhecido é transformado eletronicamente em uma lista de picos que, por meio de um algoritmo bioestatístico, é comparada com as listas de picos de microrganismos referência do banco. O valor de score correspondente é calculado a partir de três valores independentes: 1) O número de picos presentes no espectro de referência que são intimamente correlacionados com o espectro do microrganismo desconhecido é calculado, sendo que nenhuma correspondência retorna um valor =0 e uma completa correspondência retorna a um valor = 1; 2) O número de picos no espectro do microrganismo desconhecido que tem um parceiro intimamente relacionado no espectro de referência é calculado da mesma maneira que o item anterior (nenhuma correspondência = 0 e completa correspondência = 1); 3) Por último, a simetria dos pares de sinais correspondente é calculada. Desta forma, se os sinais de alta intensidade do espectro do microrganismo desconhecido corresponderem aos sinais de alta intensidade do espectro de referência do banco de dados e, o mesmo acontecer com os sinais de baixa intensidade, resultará num valor de simetria alto, com o valor da matriz de correlação próximo a 1. Se os pares correspondentes não tiverem simetria alguma, resultará em um valor próximo a 0. Vale ressaltar que o software utiliza uma abordagem superior de correspondência dos padrões, baseada na análise estatística multivariada verdadeira e inclui tanto as posições quanto intensidades e frequências dos picos, assegurando níveis altos de precisão e reprodutibilidade. Assim, esses três valores são multiplicados, o resultado é normalizado para 1000 e feito o logaritmo da pontuação. Portanto, o valor máximo obtido é 3, correspondendo ao  $\log(1000)$ . Os resultados, então, são apresentados em ordem decrescente, com a melhor correspondência no topo da lista e, quanto o maior valor do score, mais provável é a classificação ao nível de espécie (MALDI-BIOTYPER CA SYSTEM, 2014).

A simplicidade na preparação celular, sensibilidade e rapidez do método indicam que ele tem um grande potencial para a utilização de rotina na caracterização e identificação de bactérias. MALDI-TOF MS é considerada uma técnica de ionização branda gerando íons a partir de moléculas de proteínas, peptídeos e oligonucleotídeos sem fragmentação da molécula (LAY, 2000). Além disso, é relativamente resistente à presença de sais e detergentes, tornando-o adequado para a análise de amostras microbiológicas. Entretanto, a qualidade do espectro produzido é

altamente dependente da escolha da matriz, composição do solvente da matriz, processo de preparação da amostra e do método de aplicação da amostra na superfície (EVASON et al., 2001; WILLIAMS et al., 2003).

Desta forma, a iniciativa de construir um banco de dados com assinaturas espectrais dos microrganismos contaminantes de dornas de fermentação de usinas de cana-de-açúcar e de leveduras selecionadas, constitui-se um passo importante para o acervo da biodiversidade das refinarias brasileiras e uma alternativa rápida, eficiente e inovadora de identificação destes microrganismos.

### 3. JUSTIFICATIVA DO TRABALHO

Conseguir trazer técnicas inovadoras, de alta tecnologia para o setor agrícola é extremamente desafiador, pois, o uso de técnicas espectrométricas, tais como MALDI-TOF MS, LC-MS, HPLC-QTOF-MS, GC-MS, por muito tempo, foi primordialmente direcionada e aperfeiçoada em estudos clínicos, principalmente na busca de moléculas que servissem como biomarcadores para diagnóstico de doenças. A partir do sucesso em estudos clínicos e aperfeiçoamento da técnica devido ao constante desenvolvimento dos componentes constituintes do espectrômetro de massas, conferindo maior resolução e sensibilidade, percebeu-se o potencial infindável de estudar o conjunto de proteínas expressas em diferentes condições biológicas, em diversos tecidos e organismos. Desta forma, a expansão da espectrometria de massas viabilizou inúmeras possibilidades de novas áreas de atuação: estudos de estresses bióticos e abióticos em plantas, qualidade da carne, diagnóstico de doenças em animais, fatores que alteram a qualidade de bebidas fermentadas, interação microbiológica, e, agora, com este trabalho, caracterização de microrganismos presentes na indústria alcooleira.

Contudo, por mais que os estudos agrícolas se espelhem nas técnicas aprimoradas em outras áreas, existe um limite de compreensão e, conseqüentemente, necessidade de adequação das técnicas aos mecanismos moleculares e genéticos específicos dos organismos de estudo. Simplificando, é necessário que haja a máxima compreensão do objeto de estudo para o melhor direcionamento de técnicas espectrométricas e, conseqüentemente, buscar a melhor forma de responder a questão biológica em evidência. Por isso, o uso de bancos de dados espectrais preferencialmente de microrganismos clínicos, como o Biotyper, não fornece toda informação necessária para estudos em outras áreas, sendo necessário a identificação molecular para validação, e a construção de um banco de dados espectral complementar de microrganismos encontrados especificamente em ambientes fermentativos brasileiros.



#### 4. OBJETIVOS

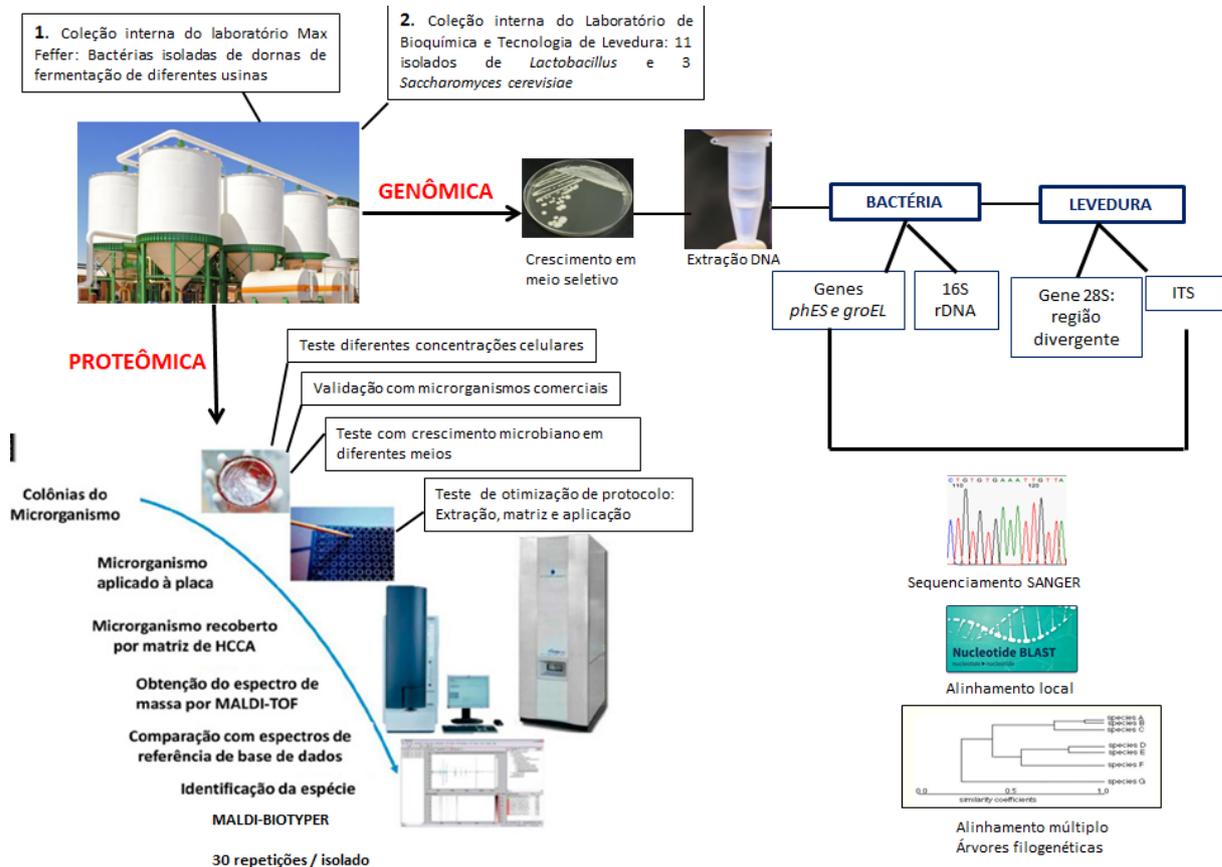
O presente trabalho é pioneiro na caracterização por espectrometria de massas de microrganismos oriundos da fermentação etanólica e objetivou:

1. Identificar as espécies de isolados bacterianos contaminantes do processo de produção de etanol, pertencentes ao gênero *Lactobacillus*, baseado no sequenciamento multigênico de regiões conservadas do genoma: 16S rDNA, *pbeS* e *groEL*;
2. Identificar três isolados de leveduras que são utilizadas como arranque no processo industrial por sequenciamento de regiões conservadas no genoma: ITS e 28S nrLSU;
3. Utilizar a identificação dos microrganismos feita por técnicas moleculares como base para a identificação do perfil protéico característico por meio de técnicas de espectrometria de massas (MALDI-TOF);
4. Validar o potencial da técnica MALDI-TOF para discriminação de diferentes espécies de isolados de *Lactobacillus*;
5. Validar o potencial da técnica para discriminação de diferentes cepas de *S. cerevisiae* selecionadas com alto desempenho no processo;
6. Construir um banco de dados espectral complementar, dentro do MALDI-Biotyper dos microrganismos fermentativos baseado na técnica de “*fingerprinting*”.



## 5. METODOLOGIA

A metodologia utilizada neste trabalho para identificação microbiana pode ser visualizada de maneira resumida no desenho experimental da Figura 4. Posteriormente, todas as etapas foram descritas separada e detalhadamente.



**Figura 4:** Desenho experimental simplificado da metodologia utilizada no projeto desde o isolamento microbiano até a identificação pela amplificação e sequenciamento de regiões específicas do genoma bacteriano e de leveduras e identificação pela obtenção de um perfil espectral proteico característico por MALDI-TOF MS.

### 5.1. Seleção dos microrganismos presentes no processo de produção de etanol

Um conjunto de microrganismos utilizados neste trabalho foram oriundos de culturas permanentes pertencentes ao Laboratório Max Feffer de Genética de Plantas (processo Fapesp: 2011/23816-0), localizado na Escola Superior de Agricultura Luiz de Queiroz (ESALQ- USP), Piracicaba, armazenadas no ultrafreezer (-80°C) em meio de cultura de Man, Rogosa, e Sharpe (1960) – MRS com adição de uma solução de glicerol na concentração de 66,7%. Os microrganismos foram isolados por trabalhos anteriores, a partir de diferentes pontos do processo de produção de etanol em diferentes usinas do Estado de São Paulo: Ponto 1 – População microbiana decorrente da cana-de-açúcar que entra na usina; Ponto 2 – População microbiana presente no caldo de cana que é enviado para a fabricação de açúcar e produção de etanol; Ponto 3 – População microbiana resistente ao processo de clarificação do caldo e do tratamento térmico realizado no Trocador de Calor a Placas (TCP); Ponto 4 – População microbiana

proveniente da fabricação de açúcar residente no mel; Ponto 5 – População microbiana presente no mosto a ser alimentado na primeira dorna de fermentação; Ponto 6 – População microbiana presente na primeira dorna de fermentação; Ponto 7 – População microbiana presente na última dorna de fermentação e; Ponto 8 – População microbiana resistente ao processo de tratamento de fermento. Como o alicerce do trabalho tratava-se de bactérias ácido lácticas contaminantes do processo fermentativo, optou-se apenas pelas coletas realizadas especificamente nas dornas (Ponto 6, 7 e 8). A partir disso, 80 microrganismos isolados de diferentes dornas foram crescidos em meio seletivo MRS e visualizados em microscopia óptica (Olympus – BX40F4, Japão) no aumento de 40x para classificação morfológica.

O segundo conjunto, composto por 11 isolados de bactérias contaminantes e 3 isolados de levedura, são de microrganismos que pertencem à coleção microbiológica selecionadas a partir do processo de produção de etanol do Laboratório de Bioquímica e Tecnologia de Leveduras, localizado também na ESALQ, Piracicaba. Estes microrganismos foram previamente classificados bioquimicamente de acordo com os metabólitos finais produzidos durante o processo de fermentação em pequena escala. Nenhum dos microrganismos pertencentes às duas coleções citadas haviam sido identificados molecularmente ou por espectrometria de massas.

Para validar previamente se a identificação realizada por meio do perfil espectral proteico característico por MALDI-TOF MS possuiria a mesma identificação remetida pelo sequenciamento de genes específicos, foram utilizadas bactérias e leveduras pertencentes à Coleção de Culturas de Microbiologia Agrícola - CCMA, situada no Laboratório de Fisiologia e Genética de Microrganismos na Universidade Federal de Lavras, que foram previamente identificadas molecularmente pelo sequenciamento parcial de genes do DNA.

## **5.2. Identificação de bactérias ácido lácticas**

### **5.2.1. Extração do DNA genômico total bacteriano**

Os isolados bacterianos foram cultivados em meio MRS por 24h a 32 °C, sob agitação (150 rpm). O protocolo de extração utilizado foi descrito por Stirling (2003) em que a cultura com crescimento saturado em tubos de 1,5 mL foi centrifugada a 10.000 rpm durante 2 minutos, até compactar o pelet. Após o sobrenadante ser descartado, o pelet foi ressuspensionado em 567µL de solução TE (Tris-HCl (1M, pH: 8,0) / 20 mL de EDTA (0,5 M, pH: 8,0) por meio de repetidas pipetagens. Em seguida, adicionou-se 30 µL de SDS (Dodecil sulfato de sódio) a 10% (concentração final de 0,5%) e 3 µL de proteinase K 20 mg/mL (concentração final de 100 µg/mL) e as amostras foram incubadas 37 °C por 1 hora. Consequente, adicionou-se 100 µL de NaCl 5M e 80 µL de CTAB (Brometo catiônico de hexadecil trimetil amônio) a 10% , incubando-as novamente a 65 ° C por 10 minutos. Após, adicionou-se 750 µL de fenol equilibrado, misturados em vortex por 40 segundos e centrifugados a 10.000 rpm durante 5 minutos. A fase superior foi removida para outro microtubo de 1,5 mL sem mexer com a interface e foi adicionado igual volume de clorofórmio: álcool isoamílico (24:1, respectivamente). Vagarosamente foram misturados e centrifugados a 10.000 rpm durante 5 minutos. Novamente, transferiu-se o sobrenadante para outro microtubo de 1.5 ml e foram adicionados 2 volumes de etanol 100% de alta qualidade. Os tubos foram acondicionados a - 20°C por 2 horas para precipitação do DNA e, posteriormente, centrifugados a 12.000 rpm por 5 minutos em temperatura ambiente. O sobrenadante foi descartado, o pelet contendo o DNA precipitado foi lavado com 600 µL de etanol a 70% e centrifugado novamente a 12.000 rpm por 5 minutos. Então, o sobrenadante foi cuidadosamente removido e lavado com 300 µL de etanol 100% e centrifugado novamente a 12.000 rpm por mais 5 minutos em temperatura

ambiente. Finalmente, o pellet foi seco e redissolvido em 50  $\mu\text{L}$  de 1/10 de solução TE + 1  $\mu\text{L}$  de RNase (1 $\mu\text{g}/\text{mL}$ ) e armazenado a  $-20^\circ\text{C}$ .

### 5.2.2. Amplificação da região 16S do rDNA

A amplificação da região V4 do gene 16S rDNA (correspondendo as posições 8-1541 no sistema numeral da *Escherichia coli*) foi realizada diretamente do DNA bacteriano extraído. Foram utilizados dois primers conservados: o reverse R1387 (5'-CGGTTACCTTGTACGACTT-3') e o forward P027F (5'-AGAGTTTGATYMTGGCTCAG-3') os quais flanqueiam as extremidades do gene de aproximadamente 1500 pb.

A amplificação foi de acordo com as seguintes condições de reação: 25 ng de DNA em 50  $\mu\text{L}$  de reação contendo 3,75 mM de  $\text{MgCl}_2$ , 0,2mM de cada dNTP, 0,2 M de cada primer, 2,5U de Taq DNA polimerase, tampão 10X. O gene 16S rDNA foi amplificado por PCR com desnaturação inicial de 10 minutos a  $94^\circ\text{C}$ , seguido de 35 ciclos de 1 minuto a  $94^\circ\text{C}$ ; 1 minuto a  $62,5^\circ\text{C}$ ; 1 minuto a  $72^\circ\text{C}$ , e uma extensão final de 10 minutos a  $72^\circ\text{C}$  em termociclador. O produto da reação foi analisado em gel de agarose (1,0 % p/v) juntamente com o marcador de massa molecular 1 Kb (Fermentas), chamado de *ladder*. Este marcador permite uma estimativa visual do tamanho de fragmentos de DNA da amostra de interesse em comparação aos fragmentos de DNA de tamanhos e concentrações conhecidas presentes no *ladder*. O *ladder* de 1kb é composto de 14 fragmentos de produtos de PCR e plasmídeos digeridos com enzimas de restrição, sendo a banda de 1000 pb a mais evidente. Assim, 1  $\mu\text{L}$  das amostras foram visualizadas em gel de agarose juntamente com o marcador molecular que abrange fragmentos de 500 pb a 10 kb.

### 5.2.3. Purificação do produto de PCR do gene 16S do rDNA

A partir da constatação de um fragmento único amplificado, os fragmentos dos genes 16S rDNA de 1500 pb foram purificados por meio da solução PEG 8000 (10g PEG 8000; 7,3 g NaCl; 45 mL água deionizada). Para tanto, 50  $\mu\text{L}$  de solução PEG foram misturados com 50  $\mu\text{L}$  do produto de PCR e incubados a  $37^\circ\text{C}$  por 15 min. Posteriormente, centrifugou-se a 10.000 g durante 15 min. O sobrenadante foi retirado com pipeta e fez-se uma lavagem primária com 125  $\mu\text{L}$  de etanol 80% gelado de alta qualidade, incubando-se durante um minuto à temperatura ambiente e então centrifugando por 5 min a 10.000 g. Retirou-se novamente o sobrenadante e repetiu-se a lavagem com etanol 80% sob as mesmas condições de incubação. Depois de centrifugadas, as amostras foram secas por 15 min a  $37^\circ\text{C}$  no speedvac. Finalmente, as amostras foram resuspendidas em 20  $\mu\text{L}$  de água ultrapura estéril e armazenadas a  $4^\circ\text{C}$  por 12 horas (overnight). Posteriormente, verificou-se a qualidade da purificação por meio de outro gel de agarose (2,0 % p/v) juntamente com o marcador de massa molecular 1Kb (Fermentas) para verificação do tamanho de 1500pb da amostra de interesse. Já para verificação da intensidade das bandas presentes no gel, utilizou-se outro marcador, Lambda DNA/HindIII, obtido a partir de DNA de fago digerido com a enzima de restrição HindIII. Foram aplicadas no gel 1 $\mu\text{L}$ , 2 $\mu\text{L}$  e 4 $\mu\text{L}$  do marcador Lambda, representando, 25 ng, 50 ng e 100 ng, respectivamente, os quais foram utilizados para estimar a quantidade de ng de DNA presente em cada banda das amostras de interesse.

Para maior confiabilidade da quantificação do fragmento amplificado, fizeram-se análises por meio da densidade óptica (DO) em espectrofotômetro. Como o DNA absorve luz no comprimento de onda de 260 nm, utiliza-se a relação de 1 DO 260nm = 50 mg de DNA de dupla fita. Assim, para o cálculo da concentração de DNA

presente na amostra, bastou utilizar a relação de DO e quantidade de DNA dupla fita vezes o fator de diluição usado na leitura do equipamento. Os resultados obtidos pelo espectrômetro foram comparados com os resultados obtidos pelos marcadores de massa molecular no gel de agarose, e 35 ng de cada amostra de DNA foram utilizados para cada reação de sequenciamento.

#### **5.2.4. Reação de sequenciamento das amostras e identificação dos microrganismos pelo gene 16S rDNA**

Para a reação de sequenciamento dos genes 16S rDNA foi realizada uma reação com BigDye® Terminator v3.1 5 x tampão de sequenciamento, BigDye® Terminator v3.1 Cycle mix de reação de sequenciamento pronto. Foram utilizados cinco primers (5 µM/µL) para amplificar o gene 16S rDNA inteiro: A) os primers conservados utilizados na reação de PCR, reverse R1387 (5'-CGGTTACCTTGTTACGACTT-3') e forward P027F (5'-AGAGTTTGTATYMTGGCTCAG -3'); B) Três intermediários (THOMPSON et al., 2001): 16F536 (5' CAGCAGCCGCGGTAATAC 3', posição 519 a 536 pb), 16F926 (5' AACTCAAAGGAATTGACGG 3', posição 908 a 926 pb) e 16F1241 (5' GCTACACACGTGCTACAATG 3', posição 1222 a 1241 pb), sendo que cada primer foi utilizado em uma reação independente. Desta forma, o produto foi composto por diferentes fragmentos, que quando sobrepostos, formaram um contig único, que corresponderá ao gene 16S completo (aproximadamente 1500pb). As amostras foram enviadas ao Laboratório Multiusuários Centralizado – Genômica Funcional Aplicada à Agropecuária e Agroenergia – ESALQ/USP para a realização do sequenciamento (<http://genfis40.esalq.usp.br/multi/>).

Para a identificação dos microrganismos contaminantes foi necessário utilizar programas de bioinformática e estabelecer filogenia do gene 16S do rDNA. Assim, a partir da leitura dos cromatogramas resultantes do sequenciamento foram obtidas as sequências de nucleotídeos correspondentes com um valor de qualidade a cada pico cromatográfico pelo programa Phred (EWING e GREEN, 1998). Para melhorar a qualidade do alinhamento das sequências, fez-se a trimagem das sequências pelo programa trimseq (<http://emboss.sourceforge.net/apps/release/6.6/emboss/apps/trimseq.html>), retirando os alinhamentos com baixa qualidade, geralmente referindo-se as pontas.

Após a obtenção de alinhamentos com qualidade, foi feito a montagem das sequências de DNA pelo programa CAP3 (HUANG & MADAN, 1999), resultantes das reações de sequenciamento a partir dos três fragmentos diferentes obtidos a partir dos 3 primers, que quando sobrepostos formam contigs únicos. Para a comparação de similaridade de sequências, será feito o download de sequências do gene 16S de rDNA de diferentes espécies de bactérias do banco de dados da plataforma EzTaxon (CHUN et al., 2007).

#### **5.2.5. Amplificação dos genes housekeeping: *pheS* e *groEL***

Os primers utilizados para a amplificação e sequenciamento do gene *pheS* foram retirados do trabalho de Naser e colaboradores, (2005<sup>a</sup>) e os primers para o gene *groEL* foram baseados no trabalho de Podlesny e colaboradores (2011), pois mostraram-se eficientes para discriminar principalmente espécies dentro do gênero *Lactobacillus*. Com o objetivo de identificar os isolados ao nível de espécie, 2 fragmentos de genes foram amplificados. Para a reação de PCR foram utilizadas 2,5 µL de tampão 10X (Sigma); 3,75 µL de MgCL<sub>2</sub> (10 mM); 1 µL de DNTP

(2,5 mM); 0,3 µL de Taq Polimerase (1un/ reação); 1 µL de cada oligonucleotídeo iniciador (10 mM), 1 µL de DNA alvo e 15,45 µL de água ultrapura estéril, totalizando o volume final da reação de 25 µL.

O modelo utilizado para a amplificação em termociclador do genes *phES* (primers 21F/23R) foi baseado em Naser e colaboradores (2005<sup>a</sup>) com adaptações: desnaturação inicial do DNA a 95°C por 10 minutos; 8 ciclos (1 minuto a 95°C, 2 minutos e 15 segundos a 57°C, 1 minuto e 15 segundos a 72°C); 18 ciclos (45 segundos a 95°C, 1 minuto e 15 segundos a 56°C, 1 minuto e 15 segundos a 72°C); 13 ciclos (45 segundos a 95°C, 1 minuto e 15 segundos a 55°C e 1 minuto 15 segundos a 72°C); extensão final de 7 minutos a 72°C. Optou-se por utilizar diferentes temperaturas de anelamento para favorecer a maior amplitude dos primers degenerados e, assim amplificar genes de diferentes espécies de lactobacilos. Já a amplificação dos genes *groEL* (primers *groEL1F/1R*) seguiu as seguintes condições de reação: desnaturação inicial a 94°C por 5 minutos, seguido de 38 ciclos de 94°C por 30 segundos, anelamento de 1 minuto a 47°C, extensão de 2 minutos a 72°C e extensão final de 10 minutos a 72°C. Os primers utilizados para a amplificação dos genes estão descritos na Tabela 1 com suas respectivas temperaturas de anelamento. Os produtos amplificados foram analisados por eletroforese em gel de agarose 1,5% e 0,5X Tris EDTA solução tampão (89 mM Tris 10x, 89 mM de ácido bórico e 25 mM de EDTA, pH 8.0) (TBE). Adicionou-se brometo de etídio (0.5 µg/ml) no gel e a visualização foi por luz U.V. Os tamanhos do fragmento esperado dos dois genes *housekeeping* é de aproximadamente 500 pb (NASER et al., 2005<sup>a</sup>; PODLESNY et al., 2011).

**Tabela 1.** Sequência dos primers utilizados para amplificação dos três genes *housekeeping* selecionados, com suas respectivas temperaturas de anelamento, posições e referência na literatura.

Primer	Sequência (5' → 3')	Gene	Temperatura de Anelamento	Posição	Referência
<i>phES- 21- F</i>	CAYCCNGCHCGYGAYATGC	<i>phES</i>	57 °C	557	Naser et al., 2005
<i>phES- 22-R</i>	CCWARVCCRAARGCAAARCC	<i>phES</i>	56 °C	1031	Naser et al., 2005
<i>groEL- 1-F</i>	GAAGGNATGAAGAAYGTBAC	<i>groEL</i>	47 °C	1	Podlesny et al., 2011
<i>groEL- 1-R</i>	AATGTHCCACGVATCTTG	<i>groEL</i>	47 °C	542	Podlesny et al., 2011

### 5.2.6. Purificação do produto de PCR do gene *housekeeping phES* e *groEL*

Como foi verificada a presença de bandas inespecíficas, toda a reação de PCR foi aplicada em gel de agarose com TAE e os fragmentos com aproximadamente 500 pb foram recortados do gel e purificados pelo Kit GFX™ PCR DNA and Gel band purification (GE Healthcare, Alemanha) de acordo com as especificações do fabricante. Posteriormente, 1 µL de cada amostra purificada foi aplicada em gel de agarose 2% TBE e comparadas com a concentração de 30 ng/ µL do marcador molecular Lambda. As amostras purificadas por kit que obtiveram uma concentração igual ou maior de 20 ng/ µL, foram separadas para o sequenciamento, já aquelas que obtiveram uma concentração inferior a 20 ng/ µL, fez-se novamente a reação de PCR e purificação por PEG- 8000 para que tivessem concentração suficiente para um sequenciamento de qualidade.

### 5.2.7. Reação de sequenciamento das amostras e identificação dos microrganismos pelos genes *housekeeping phES* e *groEL*

Para as reações de sequenciamento dos genes *housekeeping pheS* e *groEL* foram realizadas uma reação com BigDye® Terminator v3.1 5 x tampão de sequenciamento, BigDye® Terminator v3.1 Cycle mix de reação de sequenciamento pronto. Foram utilizados os mesmos pares de primers (forward e reverse) utilizados na reação de PCR descrita anteriormente (Tabela 1) na concentração de 5 µM/µL. As amostras foram enviadas ao Laboratório Multiusuários Centralizado – Genômica Funcional Aplicada à Agropecuária e Agroenergia – ESALQ/USP para a realização do sequenciamento (<http://genfis40.esalq.usp.br/multi/>).

A partir dos cromatogramas resultantes do sequenciamento, foi utilizado o Ribosomal Database Project (RDP) pipeline tool para realização da trimagem de qualidade das bases de cada sequência considerando o score de qualidade Q20 do PHRED (EWING e GREEN, 1998). Para a identificação dos isolados, as sequências de nucleotídeos serão comparadas quanto a sua similaridade com sequências de nucleotídeos presentes no banco de dados do Centro Nacional de Informação Biotecnológica dos Estados Unidos (NCBI, 2017). Foram consideradas sequências com similaridade superior ou igual a 97% para a classificação das espécies. O programa CLUSTAL X será usado para fazer o alinhamento múltiplo das sequências, o qual será importado dentro do software BioNumerics (versão 4.5) para o cálculo da matriz de similaridade (NASER et al., 2005) e construção da árvore filogenética a partir do método de agrupamento vizinhos (*Neighbor-joining methods*) (SAITOU & NEI, 1987).

### 5.2.8. Análises Filogenéticas de *Lactobacillus sp*

Para confirmar as identificações e visualizar as relações evolutivas entre os isolados de *Lactobacillus sp*, sequências dos genes *16S rRNA*, *pheS* e *groEL* juntamente com sequências do banco de dados de nucleotídeos (número de acesso nas árvores filogenéticas) foram utilizadas para as análises filogenéticas utilizando o software MEGA 6.0 (TAMURA et al., 2011), com o método Neighbor-Joining (SAITOU e NEI, 1987) e o modelo de p-distância. A divergência evolutiva também foi calculada pelo MEGA 6.0. O método maximum-likelihood (GUINDON e GASCUEL, 2003) foi utilizado para confirmar a topologia da árvore.

## 5.3. Identificação de leveduras selecionadas do processo fermentativo

### 5.3.1. Extração do DNA genômico total de eucariotos

Os isolados foram cultivados em meio Yeast extract – Peptone – Dextrose (YEPD) extrato de levedura a 1%, peptona a 2% e glicose a 2% e ágar a 2%) em placas de petri por 24 horas a 30 °C. O protocolo de extração utilizado foi descrito por Doyle e colaboradores (2003) com modificações. Assim, a cultura microbiana crescida em placa foi transferida para microtubos de 1,5 mL contendo 700 µL de solução de extração (NaCl 5M, TrisHCl 1M com pH 8, EDTA 0,5 M a pH 8, PVP-40 1%, CTAB 10%) e centrifugada a 14.000 rpm durante 7 minutos, até compactar o pelet. Após o sobrenadante ser descartado, o pelet foi liofilizado por 4 horas para secagem e triturado no moinho de bolas por 30 segundos a 30 hertz para que as células fossem rompidas. Para a extração do DNA, adicionou-se novamente 700 µL da solução de extração com adição de beta-mercaptoetanol (1%) ao material macerado, vortexou por 1 minuto e incubou por 20 minutos a 60°C. Após deixar as amostras em temperatura ambiente por 5 minutos, adicionou-se 600 µL da solução de clorofórmio: álcool isoamílico (24:1) e as amostras foram homogeneizadas vagarosamente por inversão, seguido de centrifugação a 14.000 rpm a temperatura ambiente por 7 minutos. Evitando perturbar a interfase, a fase aquosa superior foi transferida para um novo microtubo

contendo 200  $\mu\text{L}$  da solução de extração sem beta-mercaptoetanol, homogeneizado e adicionado 600  $\mu\text{L}$  da solução de clorofórmio: álcool isoamílico (24:1). Para separação das fases, as amostras foram centrifugadas a 14.000 rpm, por 7 minutos a temperatura ambiente. O sobrenadante foi novamente coletado e adicionou-se o mesmo volume coletado de isopropanol para precipitar o DNA. Vagarosamente as amostras foram homogeneizadas por inversão durante 1 minuto e centrifugadas a 14.000 rpm por 15 minutos a 4°C. Cuidadosamente, o sobrenadante foi descartado restando apenas o DNA precipitado na parte inferior do tubo. Para a lavagem do precipitado, adicionou-se 600  $\mu\text{L}$  de etanol de alta qualidade na concentração de 70%, incubou-se por 10 minutos a -20°C, seguido de centrifugação por 15 minutos a 14.000rpm a 4°C. Conseqüente, o pellet foi seco por 15 minutos a temperatura ambiente para evaporação total do etanol e ressuspendido em 45  $\mu\text{L}$  de solução tampão TE (TRIS-HCl 1M e EDTA 0.5 M com pH 8) com 10 $\mu\text{L}$ /mL de RNase. Após, as amostras foram incubadas na estufa a 37°C por 40 minutos e armazenadas por 12 horas, “overnight”, a 4°C para completa dissolução. Para verificar a qualidade do DNA, foi realizada a eletroforese com gel de agarose 1% com 5  $\mu\text{L}$  da amostra.

### 5.3.2. Amplificação da região parcial do gene ITS e 28S-nr LSU

Para a amplificação da região discriminatória nuclear ITS (internal transcribed spacer region) foram utilizados os primers ITS1 (5'-CTTGGTCATTTAGAGGAAGTAA-3') e ITS2 (5'-GCTGCGTTCTTCATCGATGC-3') descritos por Gardes e Bruns (1993) e Smith e Peay (2014). Para a amplificação do gene parcial 28S- nrLSU (*nuclear ribosomal large subunit rDNA*) foram utilizados os primers U1 (5'-GTGAAATGTGTGAAAGGGAA- 3') e U2 (5' – GACTCCTTGGTCCGTGTT- 3') descritos por Shandu et al. (1995) e Xu, Wang e Blomquist (2002). Já para a amplificação parcial do gene 28S- nrLSU correspondendo aos domínios divergentes D1/D2 ao final da subunidade ribossomal maior (nucleotídeos 63-642 em *Saccharomyces cerevisiae*) foram utilizados os pares de primers NL-1 (5'-GCATATCAATAAGCGGAGGAAAAG- 3') e NL-4 (5'- GGTCCGTGTTTCAAGACGG- 3') descritos por O'Donnell (1993). A amplificação dos três genes foi de acordo com as seguintes condições de reação: 25 ng de DNA em 50  $\mu\text{L}$  de reação contendo 3,75 mM de MgCl<sub>2</sub>, 0,2mM de cada dNTP, 0,2 M de cada primer, 2,5 U de Taq DNA polimerase, tampão 10X. O genes foram amplificados por PCR com desnaturação inicial de 10 minutos a 94 °C, seguido de 36 ciclos com desnaturação de 30 segundos a 94 °C; 1 minuto na temperatura de anelamento específica de cada primer (Tabela 2); extensão de 2 minutos a 72 °C, e uma extensão final de 7 minutos a 72 °C em termociclador.

Os produtos da reação foram analisados em gel de agarose (1,5 % p/v) com 1X tampão TE (0.04M Tris-acetate, 0.001 M EDTA [pH 8.0]), visualizados com brometo de etídio 8 x 10<sup>5</sup>  $\mu\text{g}/\mu\text{L}$ , juntamente com o marcador de massa molecular 1 Kb (Fermentas).

**Tabela 2.** Sequência dos primers utilizados para amplificação dos três genes selecionados para identificação de leveduras, com suas respectivas temperaturas de anelamento.

Primer	Sequência (5'→ 3')	Gene	Temperatura de Anelamento	Tamanho do fragmento (pb)
<i>ITS1</i>	CTTGGTCATTTAGAGGAAGTAA	ITS	52°C	600
<i>ITS2</i>	GCTGCGTTCTTCATCGATGC			
<i>U1</i>	GTGAAATTGTTGAAAGGGAA	28S	52°C	400
<i>U2</i>	GACTCCTGGTCCGTGTT			
<i>NL-1</i>	GCATATCAATAAGCGGAGGAAAAG	28S	50°C	700
<i>NL-4</i>	GGTCCGTGTTTCAAGACGG			

### 5.3.3. Purificação do produto de PCR dos genes ITS e 28S-nr LSU

As regiões ITS e 28S-nr LSU amplificadas por PCR foram purificadas utilizando a solução PEG 8000 (10g PEG 8000; 7,3 g NaCl; 45 mL água deionizada). Para isso, 50 µL de solução PEG foram adicionados a 50 µL do produto de PCR, incubou-se por 15 min a 37°C e centrifugou-se a 10.000 g durante 15 min. O sobrenadante foi retirado com pipeta e adicionou-se 125 µL de etanol 80% incubando durante 1 minuto à temperatura ambiente seguido por centrifugação a 10.000 g por 5 minutos. O sobrenadante foi novamente retirado e repetiu-se a lavagem e a incubação nas mesmas condições anteriores. Posteriormente, as amostras foram secas por 15 min a 37 °C e verificou-se a qualidade da purificação por meio de outro gel de agarose (1,5% p/v) juntamente com o marcador de massa molecular 1Kb (Fermentas) e um marcador Lambda DNA/HindIII em diferentes concentrações para estimar a intensidade das bandas aparentes no gel de agarose. Além da utilização dos marcadores moleculares, optou-se por fazer análises de densidade óptica (DO) em espectrofotômetro (Nanodrop, Thermo Scientific™) para obter uma maior confiabilidade da qualidade e quantidade de DNA a serem sequenciadas. Assim, as amostras que obtiveram quantidade de no mínimo 35ng por ambos os métodos, foram sequenciadas.

### 5.3.4. Reação de sequenciamento das amostras e identificação dos microrganismos pela região dos genes ITS e 28S-nr LSU

Para as reações de sequenciamento da região ITS e 28S-nr LSU foram realizadas uma reação com BigDye® Terminator v3.1 5 x tampão de sequenciamento, BigDye® Terminator v3.1 Cycle mix de reação de sequenciamento pronto. Foram utilizados os mesmos pares de primers (forward e reverse) utilizados na reação de PCR descrita anteriormente na concentração de 5 µM/µL. As amostras foram enviadas ao Laboratório Multiusuários Centralizado – Genômica Funcional Aplicada à Agropecuária e Agroenergia – ESALQ/USP para a realização do sequenciamento (<http://genfis40.esalq.usp.br/multi/>).

A partir dos cromatogramas resultantes do sequenciamento, foi utilizado o Ribosomal Database Project (RDP) pipeline tool para realização da trimagem de qualidade das bases de cada sequência considerando o score de qualidade Q20 do PHRED (EWING e GREEN,1998). Para a identificação dos isolados, as sequências de nucleotídeos foram comparadas quanto a sua similaridade com sequências de nucleotídeos presentes no banco de

dados do Centro Nacional de Informação Biotecnológica dos Estados Unidos (NCBI, 2017). Foram consideradas sequências com similaridade superior ou igual a 97% para a classificação das espécies. O programa CLUSTAL X foi usado para fazer o alinhamento múltiplo das sequências importadas dentro do software BioNumerics (versão 4.5) para o cálculo da matriz de similaridade (NASER et al., 2005) e construção da árvore filogenética a partir do método de agrupamento vizinhos (*Neighbor-joining methods*) (SAITOU & NEI, 1987).

### 5.3.5. Análises Filogenéticas de *Saccharomyces sp.*

Para confirmar as identificações e visualizar as relações evolutivas entre os isolados de *Saccharomyces*, sequências parciais dos genes ITS e 28S-nr LSU juntamente com sequências do banco de dados de nucleotídeos (número de acesso nas árvores filogenéticas) foram utilizadas para as análises filogenéticas utilizando o software MEGA 6.0 (TAMURA et al., 2011), com o método Neighbor-Joining (SAITOU e NEI, 1987) e o modelo de p-distância. A divergência evolutiva também foi calculada pelo MEGA 6.0. O método maximum-likelihood (GUINDON e GASCUEL, 2003) foi utilizado para confirmar a topologia da árvore.

## 5.4. Preparo de amostras para identificação dos espectros referentes aos diferentes microrganismos utilizando o instrumento analítico MALDI-TOF MS Bruker Microflex

### 5.4.1. Preparo das amostras bacterianas para aquisição por MALDI-TOF MS

Para a otimização da metodologia para a aquisição dos dados do perfil espectral proteico dos isolados bacterianos a partir das dornas das usinas foram testados alguns pontos críticos do processo de aquisição, tais como: extração proteica, modos de aplicação da amostra e escolha das matrizes.

Primeiro, em relação à extração de proteínas de bactérias ácido-láticas, foram testados quatro protocolos diferentes baseados na literatura (SENG et al., 2009; BÖHME et al., 2010; ALATOOM et al., 2011; POSTERARO et al., 2012). Os protocolos testados estão numerados e descritos abaixo:

**Teste 1** – método baseado em Böhme e colaboradores (2010): Uma alçada de células dos microrganismos selecionados crescidos previamente em meio MRS sólido por 24 horas a 30°C foi transferida para um microtubo de 1,5 mL contendo uma solução extratora 50% ACN: 1% TFA, sendo posteriormente vortexado. Consequente, as amostras foram centrifugadas a 8.000 rpm por 10 minutos a temperatura ambiente, e o sobrenadante foi transferido para um novo tubo, no qual foi estocado a -20°C até o momento das aquisições dos espectros pelo MALDI-TOF.

Para a aplicação na placa MALDI, 1 µL da amostra estocada foi transferida para um novo microtubo de 0,5 mL contendo 10 µL de matriz (Ácido  $\alpha$ -ciano-4-hidroxicinâmico, solução saturada em 70% $H_2O$ :30%ACN: 0.1%TFA) e outro teste contendo 10 µL de matriz (Ácido  $\alpha$ -ciano-4-hidroxicinâmico, solução saturada em 50% $H_2O$ :50%ACN: 2,5%TFA). Um microlitro da solução (amostra + matriz) de cada teste foi aplicada manualmente na placa MALDI e seca a temperatura ambiente. Cinco réplicas técnicas e 3 biológicas foram utilizadas para cada microrganismo testado.

**Teste 2** – método baseado em Posteraro e colaboradores (2012): Os microrganismos selecionados foram crescidos em meio MRS sólido por 24 horas a 30°C. Uma colônia isolada da placa de petri foi transferida para um

microtubo de 1,5mL e ressuspendida em 50  $\mu$ L de solução de ácido fórmico 10% e vortexado. Posteriormente, as amostras foram estocadas a  $-20^{\circ}\text{C}$  até o momento das aquisições dos espectros pelo MALDI-TOF.

Um microlitro das células lisadas foi aplicado em cada poço da placa MALDI e secas a temperatura ambiente. Logo após, os poços foram recobertos com 0.6  $\mu$ L de etanol absoluto de alta pureza e, esperou-se novamente secar em temperatura ambiente. Consequente, foram aplicados 1  $\mu$ L da solução de matriz (Ácido  $\alpha$ -ciano-4-hidroxicinâmico, solução saturada em 50% $\text{H}_2\text{O}$ :50% $\text{ACN}$ : 2,5% $\text{TFA}$ ). Cinco réplicas técnicas e 3 biológicas foram utilizadas.

**Teste 3** – método baseado em Seng e colaboradores (2009): Os microrganismos selecionados foram crescidos em meio MRS sólido por 24 horas a  $30^{\circ}\text{C}$ . Uma alçada da cultura foi transferida para um microtubo de 1.5 mL com uma solução contendo etanol absoluto e ácido fórmico 70% e vortexado. Posteriormente, as amostras foram estocadas a  $-20^{\circ}\text{C}$  até o momento das aquisições dos espectros pelo MALDI-TOF.

Um microlitro das células lisadas foi aplicado em cada poço da placa MALDI e esperou-se secar em temperatura ambiente. Logo após, aplicou-se 1  $\mu$ L da solução de matriz (ácido sinapínico em solução saturada em 50% $\text{H}_2\text{O}$ :50% $\text{ACN}$ : 2,5% $\text{TFA}$ ) e em outro teste utilizou a matriz (Ácido  $\alpha$ -ciano-4-hidroxicinâmico, solução saturada em 50% $\text{H}_2\text{O}$ :50% $\text{ACN}$ : 2,5% $\text{TFA}$ ). Cinco réplicas técnicas e 3 biológicas foram utilizadas para cada teste.

**Teste 4** – método baseado em Alatoon et al., (2011): Os microrganismos selecionados foram crescidos em meio MRS sólido por 24h a  $30^{\circ}\text{C}$ . Uma alçada de cultura foi transferida para um microtubo 1.5 mL contendo 300  $\mu$ L de água deionizada e vortexados. Posteriormente, 900  $\mu$ L de etanol absoluto foram adicionados, levemente homogeneizado e centrifugados em alta velocidade (13.000 rpm) por 2 minutos em temperatura ambiente. O sobrenadante foi descartado e o *pellet* seco no speed vac (MARCA) por 10 minutos. Na sequencia, foi adicionado sobre o *pellet* 100  $\mu$ L de ácido fórmico 70% e 100  $\mu$ L de  $\text{ACN}$  100%, e novamente agitados vagarosamente. Centrifugou-se a solução por 2 minutos a 13.000 rpm em temperatura ambiente e o sobrenadante contendo as proteínas extraídas foi coletado. As amostras foram estocadas a  $-20^{\circ}\text{C}$  até o momento das aquisições dos espectros pelo MALDI-TOF.

Um microlitro das células lisadas foi aplicado em cada poço da placa MALDI e esperou-se secar em temperatura ambiente. Logo após, aplicou-se 1  $\mu$ L da solução de matriz (Ácido  $\alpha$ -ciano-4-hidroxicinâmico, solução saturada de 10 mg/ mL em 70% $\text{H}_2\text{O}$ :30% $\text{ACN}$ : 0.1% $\text{TFA}$ ). Outros dois testes foram realizados em relação a solução da matriz: *i*) matriz ácido  $\alpha$ -ciano-4-hidroxicinâmico, solução saturada de 20 mg/ mL em 70% $\text{H}_2\text{O}$ :30% $\text{ACN}$ : 0.1% $\text{TFA}$ ) e *ii*) Matriz ácido  $\alpha$ -ciano-4-hidroxicinâmico, solução saturada de 10 mg/mL em 50% $\text{H}_2\text{O}$ :50% $\text{ACN}$ : 2,5% $\text{TFA}$ . Cinco réplicas técnicas e 3 biológicas foram utilizadas para cada teste.

Posteriormente aos testes de extração proteica e escolha de matrizes, foram testadas concentrações celulares iniciais suficientes para a obtenção de um perfil espectral de qualidade. Assim, com o auxílio de uma haste descartável estéril medidora (LabPlast) de 1  $\mu$ g, coletou-se 1, 2 e 4  $\mu$ g de células de uma mesma cultura microbiológica crescida anteriormente em placa de petri por 24h a  $30^{\circ}\text{C}$  contendo meio MRS sólido e realizou-se a extração proteica com o método de extração de melhor performance, armazenando a  $-20^{\circ}\text{C}$  para as análises de MALDI-TOF MS.

#### **5.4.2. Preparo das amostras de leveduras para aquisição por MALDI-TOF MS**

Para a extração proteica de isolados de levedura foram testados o protocolo baseado em Posteraro e colaboradores (2012), previamente descrito no item anterior (Teste-2) e o protocolo de baseado em Lauterbach e

colaboradores (2017), utilizado para microrganismos eucariotos. Assim, 1 $\mu$ g de células crescidas previamente em meio YEPD sólido por 24 horas a 30°C foram coletadas com auxílio de uma haste descartável estéril medidora e colocadas em um microtubo de 1,5 mL contendo 300  $\mu$ L de água ultrapura e vortexado por 5 minutos. Em seguida, adicionou-se 900  $\mu$ L de etanol absoluto de alta qualidade e a solução foi agitada novamente por 5 minutos. Centrifugou-se a 13.000 rpm em temperatura ambiente por 5 minutos e o sobrenadante foi descartado. Para retirar o remanescente de etanol, o *pellet* foi seco no speed vac por 20 minutos. Para a extração de proteínas, adicionou-se ao *pellet* 50  $\mu$ L de solução de ácido fórmico 70% e vortexou por 5 minutos. As amostras foram mantidas em -20°C até as aquisições por MALDI-TOF MS.

## 5.5. Aquisição dos espectros e análise dos dados obtidos por MALDI-TOF MS

### 5.5.1. Perfil proteico espectral

O espectrômetro de massas utilizado foi MALDI-TOF MS com sistema Bruker Autoflex III equipado com um feixe de laser de 334 nanômetros. Os perfis espectrais proteicos foram adquiridos em modo linear e positivo. Cada espectro foi adquirido manualmente com uma média de 5.000 disparos (1000 disparos em 5 regiões diferentes do poço), acumulados até a intensidade de  $10^4$  e uma faixa dinâmica de  $m/z$  2.000 a 15.000 foi utilizada. Antes das análises, a calibração externa do equipamento foi realizada com uma Mistura Calibrante de Padrão de Teste Bacteriano (BTS), cobrindo a faixa de massa entre 2.000 e 20.000 Da. O BTS é um extrato de *Escherichia coli* incluindo as proteínas adicionais RNase A e mioglobina. Um extrato de *Escherichia coli* (ATCC 25922) também foi usado como amostra de referência para verificar a calibração realizada anteriormente. Os espectros foram adquiridos por meio da ferramenta AutoExecute do software de aquisição Flexcontrol (versão 2.4; Bruker-Daltonik GmbH). Foram utilizados para cada microrganismo 10 repetições técnicas e 3 repetições biológicas, totalizando 30 espectros por isolado. A análise dos dados MALDI foi feita em três etapas distintas: (1) pré-processamento, (2) processamento e (3) análise estatística. Os espectros brutos foram pré-processados no software FlexAnalysis (Bruker-Daltonik) após a subtração da linha de base para remoção de fundo, alinhamento da escala de espectros, seleção de íons com uma relação sinal/ruído (S/R) maior que 3 e normalização de intensidades. Após a obtenção do perfil espectral representativo de cada isolado pela sobreposição de todos os espectros pelo software ClinProTools, os melhores espectros foram comparados com o banco de dados espectral Biotyper da Bruker para verificar se a identificação molecular correspondia à identificação espectrométrica.

### 5.5.2. Validação da técnica com microrganismos comerciais

Devido à inacessibilidade de uma coleção microbiológica acurada de microrganismos provindos de dornas de fermentação etanólica brasileiras para validar a eficiência e reprodutibilidade da técnica, foram comprados 11 microrganismos (10 do gênero *Lactobacillus* : *L. plantarum*, *L. fermentum*, *L. paracasei*, *L. brevis* e 1 *S. cerevisiae*) pertencentes a Coleção de Culturas da Microbiologia Agrícola (CCMA) da Universidade Federal de Lavras, Minas Gerais. Todos os microrganismos da coleção possuíam identificação molecular ao nível de espécie e foram isolados de ambientes fermentativos, tais como advindos da fermentação do fruto do cacau, leite, arroz, pequi. Desta forma, os isolados de *Lactobacillus* e *Saccharomyces* foram crescidos em meio MRS e YEPD sólidos, respectivamente, por 24h a

30 °C e 1 µg de células foram coletadas para a extração proteica. Tanto o preparo de amostra e aplicação na placa de MALDI, quanto escolha da matriz e obtenção do espectro foram realizados de maneira idêntica a descrita anteriormente para os isolados de dornas. Como o banco de dados Biotyper é composto principalmente por microrganismos clínicos, criou-se um banco de dados interno no software Biotyper com os espectros obtidos dos microrganismos agrícolas do CCMA, o qual foi utilizado como base para a identificação dos isolados deste estudo.

### 5.5.3. Perfil espectral em diferentes meios de cultura

A ampla capacidade de identificação bacteriana ao nível de espécie em culturas puras ou misturas microbianas simples sob condições controladas pela técnica MALDI-TOF MS já foi estabelecida por Wahl e colaboradores (2002), juntamente com o desenvolvimento de um algoritmo automatizado baseado estatisticamente para análise dos dados espectrais gerados. Entretanto, alguns trabalhos questionaram o quão eficiente seria ainda a técnica se os microrganismos fossem crescidos em diferentes meios de cultura, já que expressariam diferentes proteínas que seriam posteriormente ionizadas. Desta forma, três microrganismos comerciais (0359, 0489 e 0516) foram crescidos em quatro diferentes meios de cultura: MRS sólido; MRS líquido; MRS D-sorbitol sólido, YEPD sólido. Foram utilizadas condições de crescimento e preparo de amostra iguais as descritas anteriormente, com exceção do meio MRS líquido que teve que ser submetido a dois passos adicionais prévios, uma centrifugação e uma lavagem com água ultra pura para obtenção do pellet celular para coleta de 1 µg de células para a extração. A aquisição dos espectros e processamento dos dados por MALDI-TOF foi realizada da mesma maneira descrita anteriormente.

### 5.5.4. Regiões espectrais características de cada espécie

Após a obtenção do espectro característico de cada isolado, identificação pelo MALDI-Biotyper e identificação ao nível de espécie pelo sequenciamento molecular os isolados pertencentes à mesma espécie foram agrupados e, desta forma, foram identificadas regiões do espectro que são constantes em todos os isolados pela ferramenta Gel View do software ClinProTools (versão 3.3, Bruker).

### 5.5.5. Análise estatística multivariada a partir dos dados de MALDI-TOF MS

Para verificar se era possível diferenciar estatisticamente os microrganismos em diferentes grupos, realizaram-se análises multivariadas a partir dos perfis espectrais utilizando o software online *MetaboAnalyst* (versão 4.0, CHONG et al., 2018). Para a matriz de dados utilizou-se a lista dos picos ionizados ( $m/z$ ) e suas respectivas intensidades relativas obtidas pelo software *FlexAnalysis* (Bruker,) de cada repetição, técnica e biológica, adquirida. Os dados foram tratados antes das análises estatísticas usando scripts no R que: 1) entre as repetições técnicas, quando o  $m/z$  possuía intensidade relativa igual a 0 (zero) em mais da metade das repetições, todos os valores eram zerados, entretanto, quando possuía valor 0 (zero) em menos da metade das repetições, preenchia-se com a mediana dos valores adquiridos; 2) já entre as repetições biológicas, os zeros foram mantidos e não usou-se do cálculo da mediana para nenhum preenchimento; 3) os íons foram alinhados com uma tolerância de 10 Daltons por  $m/z$ ; 4) Apenas os

$m/z$  que possuíam valores diferentes de zero em todas as repetições, técnicas e biológicas, foram mantidos para a construção da matriz de dados. Ou seja, foram mantidos apenas os íons que tiveram 100% de frequência.

O tratamento dos dados realizado no software *MetaboAnalyst* (versão 4.0) foram baseados em Santos e colaboradores (2016). Os dados foram filtrados usando o desvio padrão para remover variáveis próximas da linha de base ou no limite de detecção. A normalização pela soma foi utilizada para ajustar as diferenças entre as amostras e autoescalamento (média centrada e dividida pelo desvio padrão de cada variável) para tornar as características individuais mais comparáveis. Ambas as abordagens estatísticas não supervisionadas e supervisionadas foram aplicadas. Análise de componentes principais (PCA), análise de mínimos quadrados parciais discriminantes (PLS-DA), análise hierárquica de agrupamentos (HCA) e importância da variável na projeção (VIP) foram executados nos dados usando o *MetaboAnalyst* (versão 4.0) para discriminar os isolados.

### **5.5.6. Teste preliminar de identificação de leveduras a partir de uma amostra de dorna industrial**

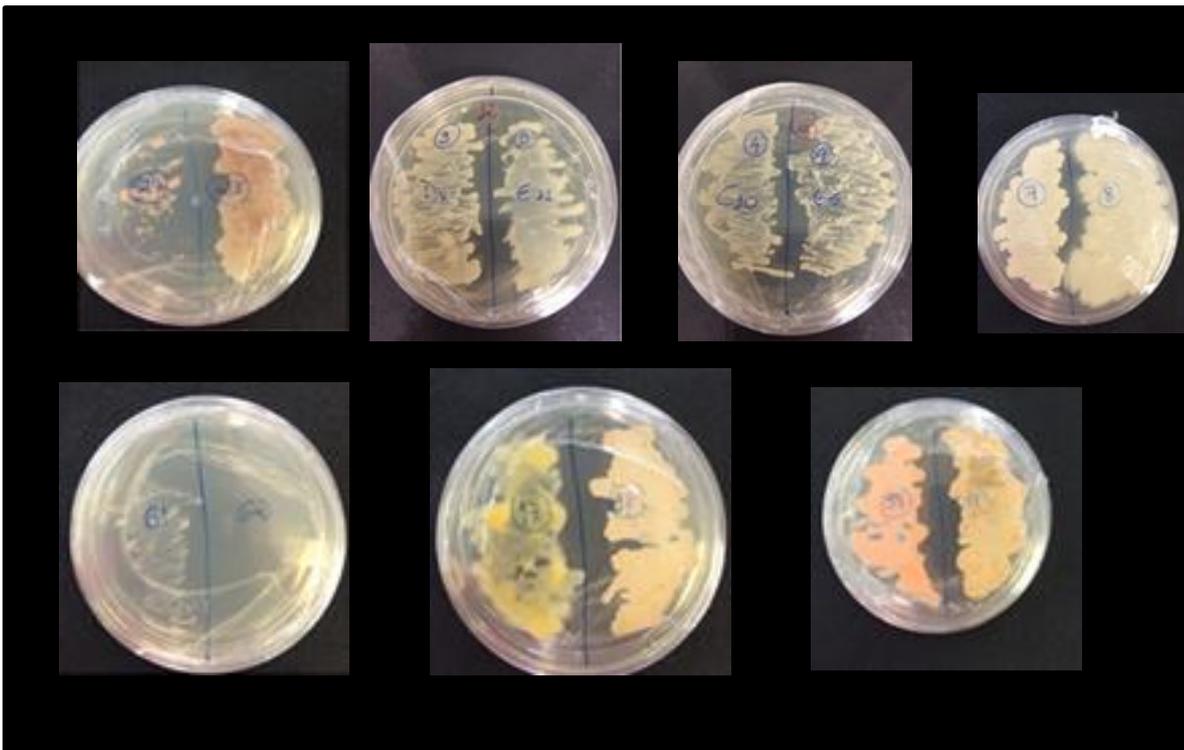
Para investigar a capacidade da técnica de MALDI-TOF MS para biotipagem de cepas industriais de *S.cerevisiae* e não-*Saccharomyces*, uma amostra de 1L diretamente da dorna em escala industrial e em condições ótimas de fermentação (contaminação controlada, menor que  $10^6$ , sem episódios de floculação e alta produtividade média – dados fornecidos pela usina) foi coletada e feito diluição seriada ( $10^5$  a  $10^8$ ) em placas de petri contento meio YEPD e antibiótico, crescidos por 24h. Oito colônias isoladas foram coletadas e preparadas para aquisição espectral por MALDI-TOF MS. Os dados espectrais brutos foram comparados com todo acervo do banco de dados MALDI-Biotyper (7.311 espectros), entre eles, espectros de *Saccharomyces* (13), *Dekkera*, (77) *Pichia*, e outros microrganismos geralmente encontrados nas dornas, considerados leveduras contaminantes do processo, juntamente com os espectros dos quatro isolados deste estudo (PEDRA- 2, SANTA-ADÉLIA, FLEISHMANN e 0681), pertencente ao banco interno complementar.

## 6. RESULTADOS E DISCUSSÃO

### 6.1. Identificação dos microrganismos por técnicas moleculares

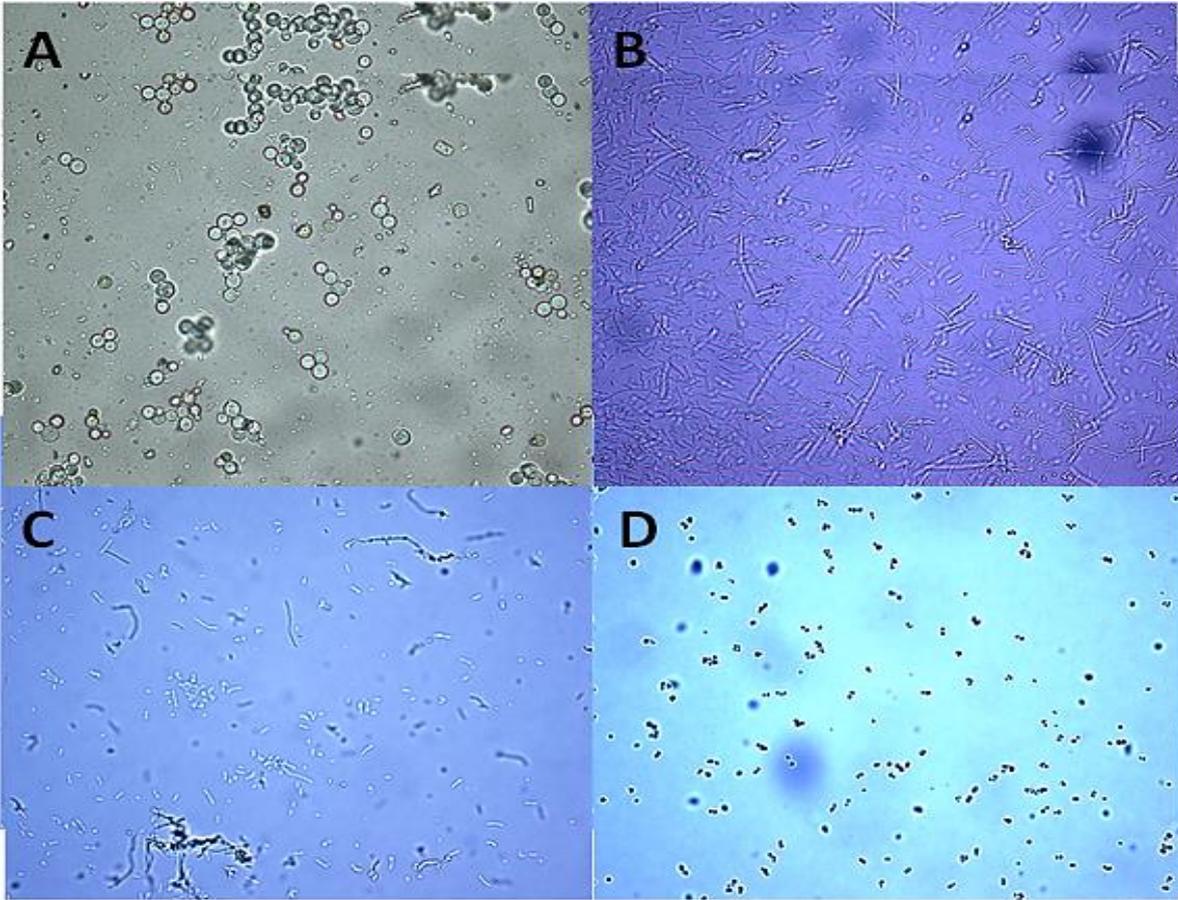
#### 6.1.1. Crescimento em meio seletivo e seleção por microscopia óptica

Os isolados utilizados neste trabalho, pertencentes tanto à coleção interna do Laboratório Max Feffer de Genética de Plantas, quanto ao Laboratório de Bioquímica e Tecnologia de Leveduras, foram crescidos em meio seletivo MRS específico para bactérias ácido lácticas e em meio YEPD para leveduras (Figura 5)



**Figura 5.** Variedade de microrganismos isolados de diferentes dornas do processo industrial de fermentação a partir de caldo de cana-de-açúcar para produção de etanol de primeira geração. Os microrganismos foram crescidos em meio MRS, a 30°C por 48h em estufa.

Com o crescimento em meio MRS, notou-se que mais de 60% dos microrganismos eram capazes de crescer no meio seletivo para lactobacilos, com uma ampla variedade de formas e cores, não sendo uma maneira eficaz de selecionar bactérias ácido lácticas quando não se usava fungicida. Como a partir de características morfológicas visíveis não foi possível caracterizar os microrganismos, optou-se por visualizar a morfologia celular em microscópio óptico, já que bactérias ácido lácticas são microrganismos que apresentam características morfológicas de cocos e bastonetes, diferenciando de outros microrganismos. Assim, uma colônia isolada, quando possível, ou uma parte pequena da colônia crescida em placa contendo meio MRS sólido de cada microrganismo foi observada em microscópio óptico no aumento de 40x (Figura 6 A - D).



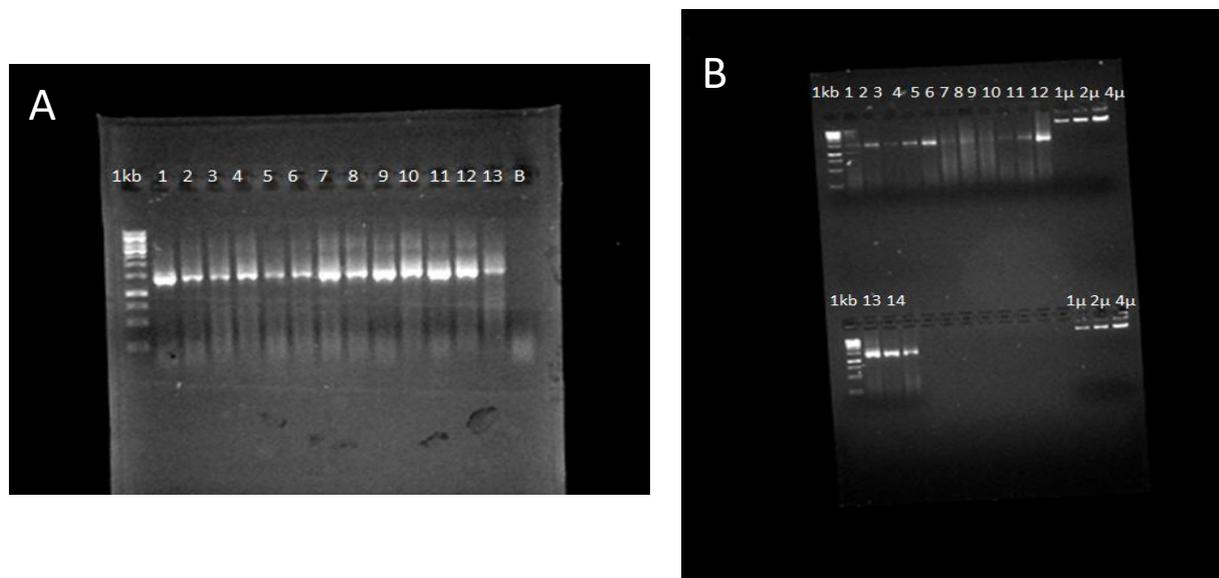
**Figura 6.** Diferentes microrganismos isolados de dornas de fermentação crescidos em meio MRS a 30°C por 48h em estufa visualizados em microscópio óptico no aumento de 40x. A) Levedura observada por microscopia óptica (células grandes, arredondadas). B e C) Bactérias do tipo bastonetes observadas por microscopia óptica, possivelmente pertencentes ao gênero *Lactobacillus*. D) Bactérias pequenas, arredondadas, do tipo coccus.

Para uma triagem de bactérias ácido lácticas, foram selecionados os microrganismos que apresentaram formas de bastonetes em microscopia óptica, possivelmente pertencentes ao gênero *Lactobacillus*. Este direcionamento foi embasado nos resultados do trabalho de Bonatelli e colaboradores (2017), o qual caracterizou a comunidade microbiana contaminante presentes em dornas de fermentação em processos industriais por meio de técnicas de metagenômica. Os resultados metagenômicos mostraram que mais de 90% dos microrganismos contaminantes presentes nas dornas pertenciam ao gênero *Lactobacillus*. Desta forma, justifica-se a especificação deste trabalho na identificação do principal gênero de bactérias contaminantes por técnicas moleculares e espectrométricas por três motivos: 1. Comprovação que esse grupo de bactérias está presente quase que em totalidade nas dornas de fermentação; 2. Principal grupo responsável pelas quedas de produção de etanol nas usinas; 3. Por se tratar de um trabalho pioneiro na elaboração de um banco de dados, definiu-se como prioridade a caracterização primária dos principais contaminantes do processo.

### 6.1.2. Amplificação da região 16S rDNA de bactérias ácido lácticas

O resultado da amplificação de aproximadamente 1500 pb da região 16S rDNA realizada diretamente do DNA extraído utilizando os primers conservados, R1387 e P027F, pode ser visualizado na Figura 7-A. Já o produto de PCR purificado com PEG-8000, pode ser visualizado na Figura 7-B. Somente os produtos de PCR purificados

que apresentaram no mínimo concentração de 35ng de DNA/  $\mu\text{L}$  é que foram utilizados para a reação de sequenciamento.

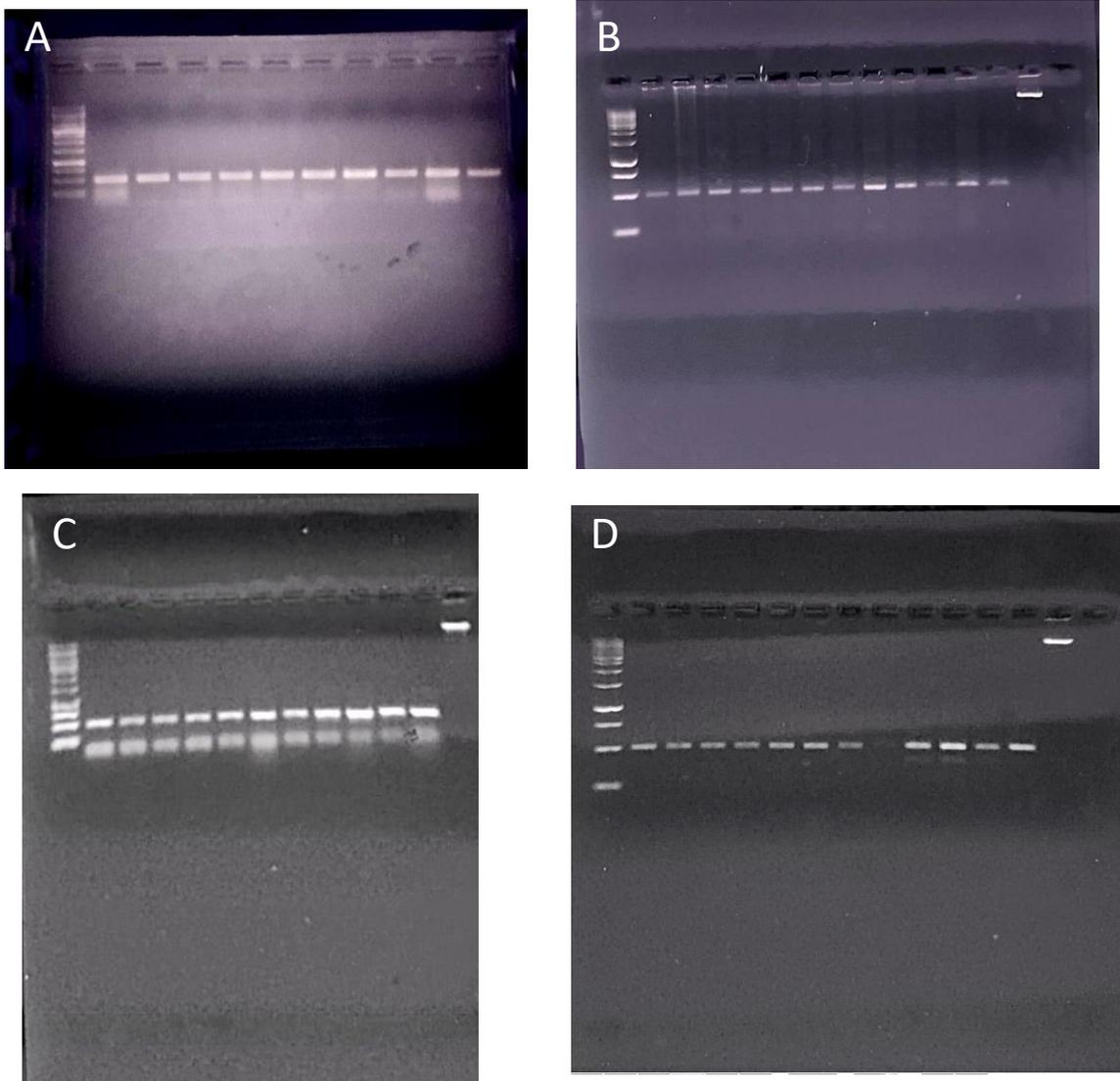


**Figura 7.** A) Produto de PCR da região 16S rDNA de aproximadamente 1500 pb de 13 bactérias isoladas do processo fermentativo etanólico (nomeadas de 1 a 13) e um branco da reação de PCR (nomeado de B), como referência foi utilizado o marcador molecular de 1kb (nomeado de 1 kb); B) Produto de PCR da região 16S rDNA de aproximadamente 1500pb purificado com PEG 8000, utilizando como referência o marcador de 1 kb (nomeado de 1 kb) e o marcador lambda aplicado em três quantidades diferentes: 1, 2 e 4 $\mu\text{L}$ , representando, 25, 50 e 100ng de DNA, respectivamente.

### 6.1.3. Amplificação dos genes housekeeping das bactérias pertencentes ao gênero *Lactobacillus*

Os resultados da amplificação de aproximadamente 500pb e purificação dos genes housekeeping *pheS* e *groEL* realizado diretamente do DNA extraído podem ser visualizados na Figura 8 (A - D). Para a amplificação dos genes *pheS* e *groEL* foram utilizados pares de primers degenerados para a identificação de diferentes espécies de lactobacilos. Entretanto, houve a amplificação de bandas inespecíficas, inviabilizando o sequenciamento direto do produto de PCR (Figura 8-A e 8-C, gene *pheS* e *groEL*, respectivamente). Desta forma, os fragmentos que apresentaram tamanho de aproximadamente 500 pb em relação ao marcador molecular 1Kb foram recortados e purificados pelo Illustra GFX PCR DNA and Gel Band Purification Kit (GE Healthcare). O resultado da purificação podem ser visualizados nas figuras 8-B e 8-D do genes *pheS* e *groEL*, respectivamente.

Visualizando a Figura 8, notamos que a estratégia de cortar as bandas de interesse e realizar suas purificações foi eficiente para eliminação de bandas inespecíficas que poderiam impedir um sequenciamento de qualidade e, conseqüentemente, prejudicar na identificação dos isolados. Entretanto, algumas amostras tiveram perda de DNA durante o processo de purificação, resultando em uma quantidade de DNA inferior a necessária para realizar as reações de sequenciamento (20 ng/  $\mu\text{L}$ ). Para essas amostras, realizou-se uma nova reação de PCR com o produto purificado com os primers e condições de reação já descritos, seguido por uma purificação por PEG-8000. Constatando que os produtos de PCR possuíam quantificação igual ou superior a 20 ng/  $\mu\text{L}$ , as amostras foram enviadas para sequenciamento.



**Figura 8.** Géis de agarose (2%) com brometo de etídio (0.5 µg/mL) e a visualização foi por luz U.V dos fragmentos amplificados e purificados dos genes *housekeeping*. – A) Produto de PCR do gene *pbeS* de aproximadamente 500 pb de bactérias isoladas do processo fermentativo etanólico, como referência foi utilizado o marcador molecular de 1kb (nomeado de 1 kb); B) Produto de PCR purificado do gene *pbeS*, utilizando como referência o marcador de 1kb (nomeado de 1 kb) e o marcador lambda na concentração de 25 ng/µL; C) Produto de PCR do gene *groEL* de aproximadamente 500 pb de bactérias isoladas do processo fermentativo etanólico, como referência foi utilizado o marcador molecular de 1kb (nomeado de 1 kb); D) Produto de PCR purificado do gene *groEL*, utilizando como referência o marcador de 1kb (nomeado de 1 kb) e o marcador lambda na concentração de 25 ng/µL.

#### 6.1.4. Identificação molecular a nível de espécie de isolados de *Lactobacillus* por sequenciamento de regiões gênicas conservadas: 16S rDNA , *pheS* e *groEL*

As identificações dos microrganismos foram realizados a partir do sequenciamento parcial dos genes 16S rDNA, *pheS* e *groEL* e alinhamento local pela ferramenta BLASTn (*Nucleotide Basic Local Alignment Search Tool*) contra o banco de dados público do NCBI (Nacional Center of Biotechnology Information). Os alinhamentos locais permitem a busca por sequências similares em bancos de dados a uma sequência específica de interesse, em que visualiza quão bem duas sequências se alinham a cada resíduo comparado, identificando os respectivos *bits* com

diferentes *scores* e valores de *e-value*. O BLAST é a ferramenta de alinhamento local mais utilizada no mundo e segue o modelo de alinhamento simples, local e heurístico e para cada *hit* é associada uma pontuação (*score*) que quantifica o grau de similaridade entre as duas sequências alinhadas (CLAVERIE e NOTREDAME, 2007). Outro ponto importante sobre os alinhamentos resultantes do BLAST é que eles representam sequências similares e não necessariamente em sequências homólogas, ou seja, mesmo considerando um alto *score* entre duas sequências, nem sempre é possível inferir um alto grau de homologia entre elas devido à aleatoriedade no processo de alinhamento, podendo ser sequências que não tenham relacionamento evolutivo algum, principalmente quando se compara organismos muito distantes na escala evolutiva. Devido a esse entrave, a estatística do programa é baseada em um valor chamado expected-value (*e-value*), o qual fornece a segurança que um *hit* não ocorreu de maneira aleatória, assim quanto mais próximo de zero for o valor do *e-value*, mais confiável será o alinhamento. Dentre as dezenas de *hits* resultantes do BLASTn, foram selecionados como melhores alinhamentos aqueles que possuíam uma cobertura de 100%, *score* máximo e *e-value* com valor de zero, quando possível. Os resultados das identificações moleculares podem ser visualizados na Tabela 3.

Como podemos notar, o gene 16S rDNA sozinho não foi suficientemente discriminatório para identificar ao nível de espécie para 72% dos microrganismos, mesmo quando havia 100% de cobertura e 99% de similaridade entre a sequência *query* (de interesse) e as sequências encontradas nos bancos de dados (*subject*), sendo apenas discriminatório para espécies de *L. plantarum*. Estudos mostram que o gene 16S rRNA evolui lentamente, além de não possuir sítios diagnósticos suficientes e a homologia da sequência 16S rDNA varia de 90,9% a 99% entre várias espécies de *Lactobacillus*, o que torna muitas vezes a discriminação ambígua (SINGH et al., 2009). Enquanto que, os genes *pheS* e o *groEL* tiveram um poder discriminatório maior, possibilitando identificar um número maior de isolados ao nível de espécie. Esses resultados corroboram com os resultados de outros trabalhos (NASER et al., 2005; NASER et al., 2007; BRUYNE et al., 2008; MADOROBA et al., 2011; OKI et al., 2011; NGUYEN et al., 2013; e PODLESNY et al., 2011), que mostraram que o sequenciamento de genes codificadores de proteínas é uma alternativa de identificação eficaz, possuindo um poder discriminatório maior em relação ao gene 16SrRNA na discriminação entre espécies de LAB, entretanto, para isolados altamente relacionados como *L. casei* e *L. paracasei*, o gene *groEL* não foi muito discriminatório.

**Tabela 3.** Identificação dos microrganismos por meio do alinhamento local realizado no programa Basic Local Alignment Search Tool (BLAST) que foram sequenciados a partir da amplificação do gene 16S rDNA, *phES* e *groEL*.

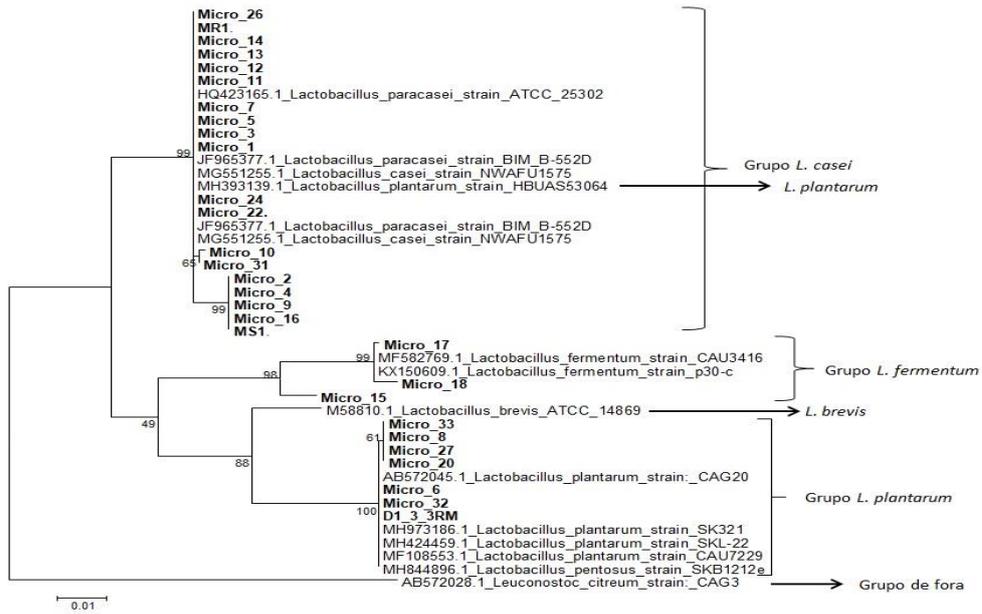
Código	Gene amplificado	Cobertura	E-value	Similaridade	Número do acesso	Id. Molecular (NCBI)
MICRO1	PHES	98%	3.00E-176	97%	<a href="#">KY502278.1</a>	<i>L.paracasei</i>
	GROEL	100%	0	100%	<a href="#">JF520448.1/FJ984016.1</a>	<i>L.paracasei/L.casei</i>
	16S	99%	0	99%	<a href="#">MK026811.1</a>	<i>Lactobacillus sp</i>
MICRO2	PHES	100%	0	100%	<a href="#">KY502278.1</a>	<i>L.paracasei</i>
	GROEL	100%	0	100%	<a href="#">JF520448.1/FJ984016.1</a>	<i>L.paracasei/L.casei</i>
	16S	99%	0	96%	<a href="#">KC836607.1</a>	<i>Lactobacillus sp</i>
MICRO3	PHES	100%	0	100%	<a href="#">KY502278.1</a>	<i>L.paracasei</i>
	GROEL	100%	0	100%	<a href="#">JF520448.1/FJ984016.1</a>	<i>L.paracasei/L.casei</i>
	16S	100%	0	99%	<a href="#">MH656948.1</a>	<i>Lactobacillus sp</i>
MICRO4	PHES	100%	0	100%	<a href="#">KY502278.1</a>	<i>L.paracasei</i>
	GROEL	100%	0	100%	<a href="#">JF520448.1/FJ984016.1</a>	<i>L.paracasei/L.casei</i>
	16S	99%	0	96%	<a href="#">KC836607.1</a>	<i>Lactobacillus sp</i>
MICRO5	PHES	100%	2.00E-168	100%	<a href="#">KY502278.1</a>	<i>L.paracasei</i>
	GROEL	100%	0	100%	<a href="#">JF520448.1/FJ984016.1</a>	<i>L.paracasei/L.casei</i>
	16S	100%	0	99%	<a href="#">CP032637.1</a>	<i>Lactobacillus sp</i>
MICRO6	PHES	100%	1.00E-139	99%	<a href="#">AM694185.1</a>	<i>L.plantarum</i>
	GROEL	100%	0	100%	<a href="#">AY424354.1</a>	<i>L.plantarum</i>
	16S	100%	0	99%	<a href="#">MF429747.1</a>	<i>L.plantarum</i>
MICRO7	PHES	100%	0	100%	<a href="#">KY502278.1</a>	<i>L.paracasei</i>
	GROEL	100%	0	100%	<a href="#">JF520448.1/FJ984016.1</a>	<i>L.paracasei/L.casei</i>
	16S	100%	0	99%	<a href="#">MK026811.1</a>	<i>Lactobacillus</i>
MICRO8	PHES	100%	2.00E-163	100%	<a href="#">AM694185.1</a>	<i>L.plantarum</i>
	GROEL	----	----	-----	----	----
	16S	100%	0	99%	<a href="#">CP023728.1</a>	<i>L.plantarum</i>
MICRO9	PHES	100%	0	100%	<a href="#">KY502278.1</a>	<i>L.paracasei</i>
	GROEL	100%	1.00E-180	100%	<a href="#">JF520448.1/FJ984016.1</a>	<i>L.paracasei/L.casei</i>
	16S	99%	0	99%	<a href="#">MH935901.1</a>	<i>Lactobacillus sp</i>
MICRO10	PHES	100%	0	100%	<a href="#">KY502278.1</a>	<i>L.paracasei</i>
	GROEL	100%	0	100%	<a href="#">KX808516.1/FJ984016.1</a>	<i>L.paracasei/L.casei</i>
	16S	99%	0	99%	<a href="#">MH935901.1</a>	<i>Lactobacillus sp</i>
MICRO11	PHES	100%	0	100%	<a href="#">KY502278.1</a>	<i>L.paracasei</i>
	GROEL	100%	0	100%	<a href="#">JF520448.1/FJ984016.1</a>	<i>L.paracasei/L.casei</i>
	16S	99%	0	99%	<a href="#">MH935901.1</a>	<i>Lactobacillus sp</i>
MICRO12	PHES	100%	0	100%	<a href="#">KY502278.1</a>	<i>L.paracasei</i>
	GROEL	100%	0	99%	<a href="#">FJ984016.1</a>	<i>L.casei</i>
	16S	99%	0	99%	<a href="#">JN415160.1</a>	<i>Lactobacillus sp</i>
MICRO13	PHES	100%	0	99%	<a href="#">KY502278.1</a>	<i>L.paracasei</i>
	GROEL	95%	6.00E-179	99%	<a href="#">KX808516.1/FJ984016.1</a>	<i>L.paracasei/ L.casei</i>
	16S	100%	0	99%	<a href="#">MH392818.1</a>	<i>Lactobacillus sp</i>
MICRO14	PHES	100%	2.00E-168	99%	<a href="#">KF746169.1</a>	<i>L.casei</i>
	GROEL	100%	0	100%	<a href="#">FJ984017.1</a>	<i>L.casei</i>
	16S	100%	0	99%	<a href="#">MG551255.1</a>	<i>Lactobacillus sp</i>
MICRO15	PHES	----	----	-----	----	----
	GROEL	100%	2.00E-50	90%	<a href="#">JF520448.1/FJ984016.1</a>	<i>L.paracasei/ L.casei</i>
	16S	99%	0	97%	<a href="#">KP284861.1</a>	<i>Lactobacillus sp</i>
MICRO16	PHES	100%	0	100%	<a href="#">KY502278.1</a>	<i>L.paracasei</i>
	GROEL	100%	0	100%	<a href="#">FJ983991.1</a>	<i>L.casei</i>
	16S	100%	0	99%	<a href="#">KC967212.1</a>	<i>Lactobacillus sp</i>
MICRO17	PHES	100%	3.00E-176	100%	<a href="#">AM284189.1</a>	<i>L.fermentum</i>
	GROEL	---	----	----	----	----
	16S	99%	0	99%	<a href="#">MF582785.1</a>	<i>Lactobacillus sp</i>
MICRO18	PHES	100%	1.00E-120	90%	<a href="#">KF746170.1</a>	<i>L.fermentum</i>
	GROEL	100%	0	100%	<a href="#">KJ807022.1</a>	<i>L.fermentum</i>
	16S	99%	0	99%	<a href="#">MF354521.1</a>	<i>Lactobacillus sp</i>

Código	Gene amplificado	Cobertura	E-value	Similaridade	Número do acesso	Id. Molecular (NCBI)
MICRO20	PHES	100%	0	100%	<a href="#">AM694185.1</a>	<i>L. plantarum</i>
	GROEL	100%	0	100%	<a href="#">AY424354.1</a>	<i>L. plantarum</i>
	16S	100%	5.00E-102	100%	<a href="#">KX815879.1</a>	<i>L. plantarum</i>
MICRO22	PHES	100%	7.00E-173	100%	<a href="#">MG932095.1</a>	<i>L. plantarum</i>
	GROEL	100%	0	100%	<a href="#">KJ807054.1</a>	<i>L. plantarum</i>
	16S	100%	0	99%	<a href="#">MH916634.1</a>	<i>L. plantarum</i>
MICRO24	PHES	100%	0	100%	<a href="#">AM694185.1</a>	<i>L. plantarum</i>
	GROEL	100%	0	100%	<a href="#">FJ984043.1</a>	<i>L. plantarum</i>
	16S	100%	0	99%	<a href="#">MH916634.1</a>	<i>L. plantarum</i>
MICRO26	PHES	100%	4.00E-170	100%	<a href="#">KY502278.1</a>	<i>L. paracasei</i>
	GROEL	100%	0	100%	<a href="#">JF520448.1</a> / <a href="#">FJ984016.1</a>	<i>L. paracasei</i> / <i>L. casei</i>
	16S	99%	0	99%	<a href="#">HM151340.1</a>	<i>Lactobacillus sp</i>
MICRO27	PHES	100%	2.00E-172	100%	<a href="#">MG932095.1</a>	<i>L. plantarum</i>
	GROEL	100%	0	100%	<a href="#">KJ807054.1</a>	<i>L. plantarum</i>
	16S	100%	0	99%	<a href="#">MH973186.1</a>	<i>L. plantarum</i>
MICRO31	PHES	100%	0	99%	<a href="#">KY502278.1</a>	<i>L. paracasei</i>
	GROEL	99%	2.00E-178	95%	<a href="#">JF520448.1</a> / <a href="#">FJ984036.1</a>	<i>L. paracasei</i> / <i>L. casei</i>
	16S	100%	0	99%	<a href="#">AB969779.1</a>	<i>Lactobacillus sp</i>
MICRO32	PHES	100%	0	100%	<a href="#">AM694185.1</a>	<i>L. plantarum</i>
	GROEL	100%	0	100%	<a href="#">FJ984065.1</a>	<i>L. plantarum</i>
	16S	100%	0	99%	<a href="#">MF429747.1</a>	<i>L. plantarum</i>
MICRO33	PHES	100%	0	100%	<a href="#">AM694185.1</a>	<i>L. plantarum</i>
	GROEL	100%	0	100%	<a href="#">AY424354.1</a>	<i>L. plantarum</i>
	16S	100%	0	99%	<a href="#">MH393133.1</a>	<i>L. plantarum</i>
MS1	PHES	100%	0	100%	<a href="#">KY502278.1</a>	<i>L. paracasei</i>
	GROEL	100%	0	100%	<a href="#">JF520448.1</a> / <a href="#">FJ984016.1</a>	<i>L. paracasei</i> / <i>L. casei</i>
	16S	100%	0	99%	<a href="#">KC967212.1</a>	<i>Lactobacillus sp</i>
MR1	PHES	100%	2.00E-168	100%	<a href="#">KF746169.1</a>	<i>L. casei</i>
	GROEL	100%	0	100%	<a href="#">FJ984017.1</a>	<i>L. casei</i>
	16S	100%	0	100%	<a href="#">MG551255.1</a>	<i>Lactocillus sp</i>
D133RM	PHES	100%	0	100%	<a href="#">AM694185.1</a>	<i>L. plantarum</i>
	GROEL	100%	0	100%	<a href="#">AY424354.1</a>	<i>L. plantarum</i>
	16S	100%	0	100%	<a href="#">CP022294.1</a>	<i>L. plantarum</i>

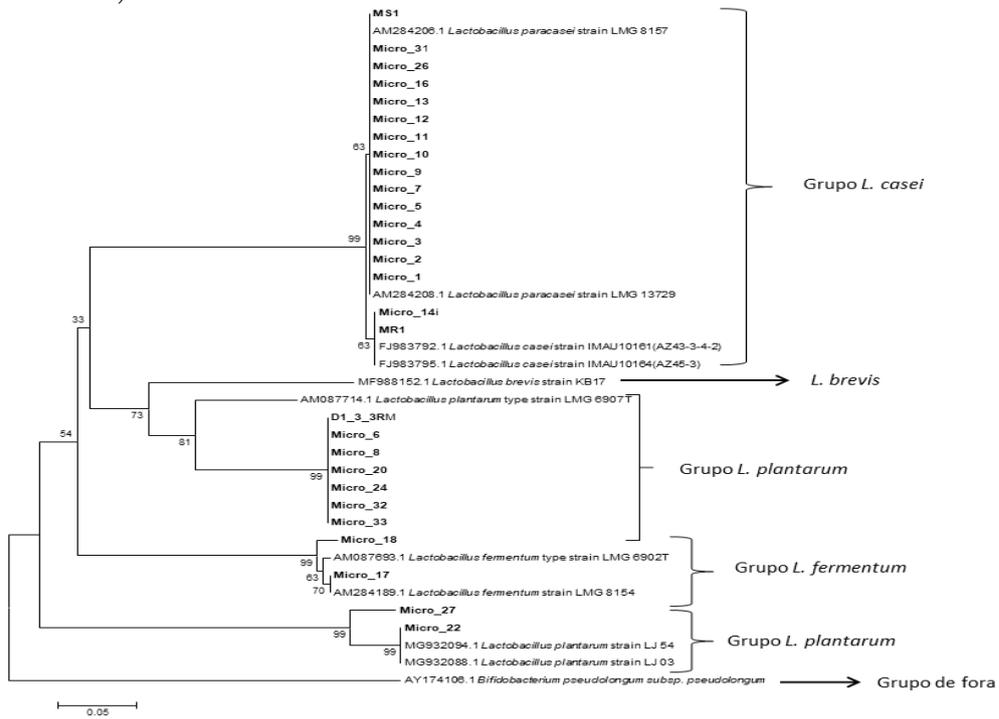
\*Código = código utilizado para identificação da amostra; 16S = referente ao gene 16S rDNA; pheS = referente ao gene housekeeping fenilalanina tRNA sintetase, subunidade alfa; groEL = referente as proteínas de choque térmico hsp60; Similaridade = quão próximo é a sequência de interesse quando comparada a sequências de banco de dados públicos do NCBI; Cobertura = porcentagem da sequência submetida ao programa que é coberta pelo alinhamento; Número de Acesso = Código de referência da sequência similar a sequência de interesse encontrada no banco de dados do NCBI. E-value = parâmetro de confiança; L. = abreviatura do gênero *Lactobacillus*; Id. = abreviatura para a palavra identificação; Id. Molecular = identificação referente ao resultado do alinhamento local no NCBI; traço (-----) = ausência de resultados

Para verificar o real poder discriminatório dos genes housekeeping em comparação com a região ribossomal bacteriana, foram construídas árvores filogenéticas. A Figura 9 (A – C) mostra as árvores filogenéticas construídas usando seqüências 16S rDNA, *phES* e *groEL*, respectivamente, dos isolados deste trabalho e cepas de referência ou de tipo relevantes recuperadas do banco de dados de nucleotídeos (<http://www.ncbi.nlm.nih.gov/blast>). As topologias das árvores filogenéticas, geradas pelo método *Neighbor-Joining*, foram confirmadas pela topologia de árvore gerada pelo método da máxima verossimilhança.

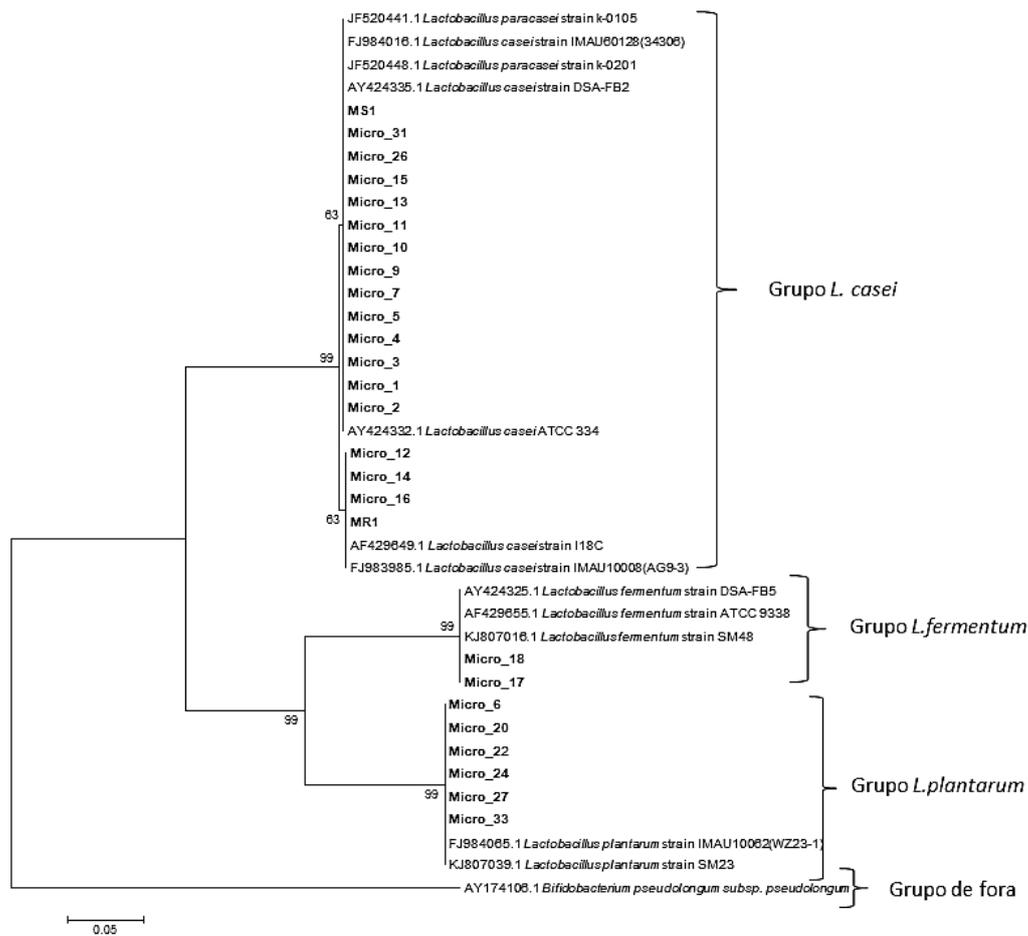
A)



B)



C)



**Figura 9.** A) Árvore filogenética baseada nas sequências do gene 16S rDNA construída usando o método neighbor-joining. *Leuconostoc citreum CAG3* foi selecionado como grupo de fora (*out group*) e 5% substituições por sítio. B) Árvore filogenética dos genes *pheS* construída usando o método neighbor-joining. *Bifidobacterium pseudolongum subsp. pseudolongum* (AY174106) foi selecionado como grupo de fora (*out group*) e 5% substituições por sítio. C) Árvore filogenética dos genes *groEL* construída usando o método neighbor-joining. *Bifidobacterium pseudolongum subsp. pseudolongum* (AY174106) é o grupo de fora (*out group*) e 5% substituições por sítio.

\*Os valores de bootstrap maiores que 60% estão mostrados nos internós. Siglas em negrito pertencem aos microrganismos deste estudo. Siglas sem negrito pertencem ao código da sequência pertencente ao banco de dados de nucleotídeos.

Avaliando a árvore filogenética criada com as sequências 16S do rDNA (Figura 9-A) notamos que o grupo denominado *L. casei* possui um grupo maior com 18 sequências, sendo 12 sequências dos isolados deste estudo e 6 pertencentes ao banco de dados; e dois subgrupos menores, de 2 e 5 sequências, respectivamente, com sequências pertencentes apenas aos isolados de dornas. Além disso, notamos que as sequências do banco de dados deste grupo pertencem a diferentes espécies: *Lactobacillus casei*, *Lactobacillus paracasei* e 1 isolado de *Lactobacillus plantarum*, não havendo distinção entre as mesmas pela região 16S do rDNA. Outros estudos (TYNKKYNEN et al., 1999; ZHONG et al., 1998; NASER et al., 2007) relataram a dificuldade de distinguir isolados pertencentes ao grupo *L. casei*, pois a similaridade na sequência do gene 16S rDNA pode ser de 99% entre os membros do grupo (MORI et al., 1997). A posição taxonômica do grupo *L. casei* (formado por espécies de *L. casei*, *L. paracasei* e *L. rhamnosus*) já foi debatida e modificada anteriormente. No Manual de Bacteriologia Sistemática lista quatro subespécies de *L. casei* (KANDLER e WEISS, 1986), os quais foram reclassificados de acordo com dados de homologia de DNA em três espécies por Collins e colaboradores (1998). Entretanto, trabalhos realizados por Klijn e colaboradores (1991) e Mori e colaboradores (1997) observaram diferenças na região V1 do 16S rRNA que permitiam a distinção das três espécies

pertencentes ao grupo *L.casei* e posteriormente, Ward e Timmis (1999) desenvolveram primers específicos para a região V1 que permitiu a distinção de 63 isolados pertencentes ao grupo *L.casei*, 51 foram classificados como *L.paracasei* e 12 como *L. rhamnosus*, não havendo nenhuma distinção entre *L.casei* e *L.paracasei* especificamente. Já Boturri e colaboradores (2017) buscando por genes que efetivamente eram capazes de discriminar isolados pertencentes ao *L.casei* grupo, testou diferentes genes housekeeping como alternativa ao gene 16S rDNA, que já se mostrou insatisfatório em muitos trabalhos e observou que a técnica MLST é realmente muito promissora.

Shevtsov e colaboradores (2010) tentando identificar diferentes espécies de *Lactobacillus*, utilizou diferentes genes *housekeeping*, entre eles o gene *groEL* e confrontou com a classificação obtida pela região 16S rDNA, notou que a porcentagem de similaridade entre as seqüências nucleotídicas dos fragmentos do gene *groEL* para *L. casei* e *L. paracasei* foi de 87,0%, enquanto que pelo gene 16S rDNA ultrapassou 98%. No entanto, pela árvore filogenética construída a partir do gene *groEL* (Figura 9-C), notamos resultados similares ao gene 16S em relação ao grupo maior pertencente ao grupo *L. casei*, com excessão que apenas cepas pertencentes as espécies *L. casei* e *L. paracasei* pertenceram ao mesmo nó. Entretanto, houve uma subdivisão de 4 isolados deste estudo (Micro\_12, Micro\_14, Micro\_16 e MR1) juntamente com as cepas AF429649.1 e FJ983985.1 ambas classificadas como *Lactobacillus casei*, mostrando que o gene *groEL* teve um poder discriminatório maior do que o gene 16S rDNA, suficiente para discriminar alguns isolados pertencentes a espécies altamente relacionadas do grupo, concordando com os resultados obtidos por Shevtsov e colaboradores (2010).

Estudos realizados por Naser e colaboradores (2007) declararam que isolados da espécie *L.casei* e *L.paracasei* foram distinguidos com base nos genes housekeeping *pheS* e *rpoA*, os quais possuíam 84% e 95%, respectivamente, de similaridade na seqüência entre os isolados do grupo. Em estudos anteriores, Naser e colaboradores (2005) analisando a aplicação de análise de seqüências de multilocus para a identificação de espécies de *Enterococcus*, notou que à nível interespecífico, o gene *pheS* mostrou um alto grau de resolução para diferenciar tais espécies, sendo que a similaridade do gene *pheS* foi no máximo de 86% para todas as espécies estudadas, entretanto algumas cepas específicas possuíam até 97% de similaridade. Esses resultados corroboram com os resultados encontrados neste estudo sobre o poder discriminatório do gene *pheS* (Figura 9-B). Analisando o grupo *L.casei* da árvore filogenética construída a partir de seqüências do gene *pheS*, notamos que o grupo foi dividido em dois subgrupos: um grupo com um nó de isolados do estudo juntamente com cepas de referência *Lactobacillus paracasei* (AM284206.1 e AM284208.1); e outro nó com dois isolados do estudo (Micro\_14 e MR1) juntamente com as cepas de referência *Lactobacillus casei* (FJ983792.1 e FJ983795.1). Desta forma, notamos a clara distinção entre isolados altamente correlacionados deste grupo, sendo que o gene *pheS* teve um alto poder discriminatório. Adicionalmente, analisando os dois genes housekeeping concomitantemente (Figura 9-B e C), temos que apenas 3 isolados (Micro\_12, Micro\_13 e Micro\_16) possuíam classificações diferentes, sendo denominados *L. paracasei* pelo gene *pheS* e *L. casei* pelo gene *groEL* e 2 isolados (Micro\_14 e MR1) foram classificados como *L. casei* por ambos os genes *housekeeping*.

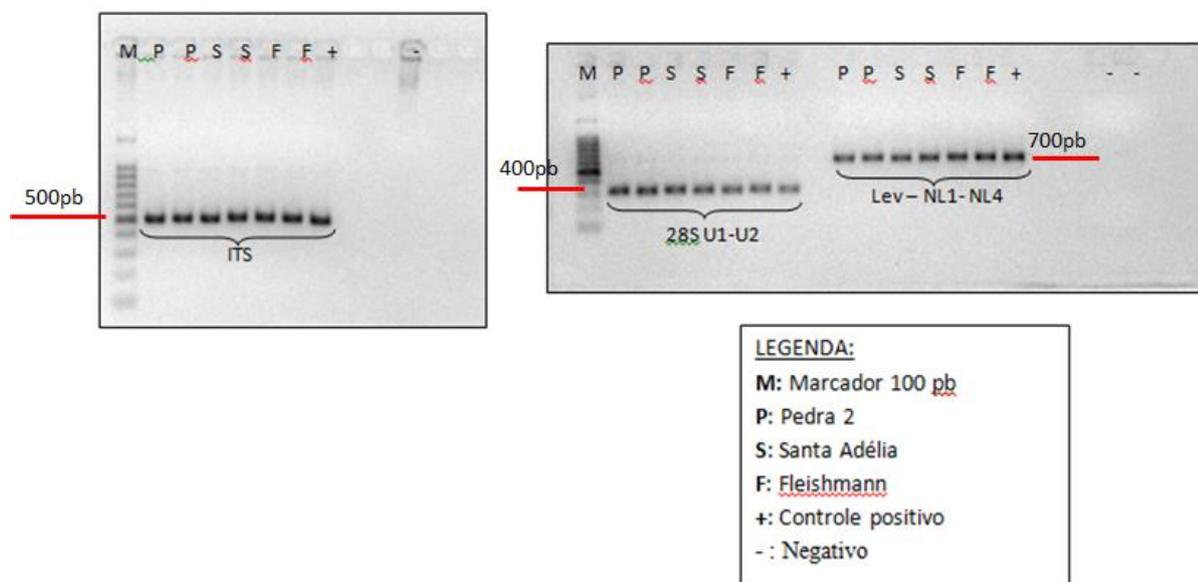
Em relação ao segundo maior grupo, dos *L. plantarum*, obtive 9 isolados pelo gene *pheS*, 6 pelo gene *groEL* e 7 pelo 16S rDNA (Figura 9 A – C). Essa diferença do número de componentes se dá pelo fato de que 3 microrganismos (Micro\_8, Micro\_32 e D1\_3\_3RM) não obtiveram qualidade no sequenciamento pelo gene *groEL*, sendo retirados da análise filogenética. Já em relação ao gene 16S rDNA, dois dos isolados (Micro\_22 e Micro\_24) classificados como *L. plantarum* por ambos genes *housekeeping*, foram classificados como pertencentes ao grupo *L. casei* pelo gene 16S. Esses resultados enfatizam mais uma vez a ambiguidade da classificação utilizando apenas a região 16S rDNA citada por Sing e colaboradores (2009), que relaciona a evolução muito lenta do gene 16S, gerando

um alto grau de similaridade entre as sequências (99%), impossibilitando a discriminação precisa entre espécies muito próximas e evidencia o maior poder discriminatório dos genes *housekeeping* também para espécies de *L. plantarum*. Avaliando ainda o grupo *L. plantarum* nas três árvores, notamos que enquanto o gene 16S rDNA e *groEL* formaram apenas um nó com todos os isolados, este grupo sofreu uma divisão pelo gene *pheS* formando dois nós diferentes com 2 subdivisões cada um.

O grupo *L. fermentum* foi composto por apenas dois isolados (Micro\_17 e Micro\_18) nas três árvores filogenéticas, entretanto a topologia da árvore não se manteve a mesma para nenhum dos três genes. Enquanto para o gene *groEL* o grupo se manteve em um único nó, para os genes *pheS* e 16S rDNA, houve a formação de subgrupos, sendo que o Micro\_18 pertencia a um ramo isolado, enquanto o Micro\_17 permaneceu no mesmo nó que as cepas de banco de dados. Interessantemente, para esse grupo não houve ambiguidade na identificação pelo gene 16S e o gene *groEL* foi o que teve menor poder discriminatório ao contrário do que ocorreu com os grupos anteriores.

### 6.1.5. Identificação de leveduras pelos genes ITS, 28S-nr LSU

Os resultados da amplificação dos 3 genes ITS (aproximadamente 600pb), 28S-U1/U2 (aproximadamente 400pb) e 28S-NL1/NL4 (aproximadamente 700pb) dos três isolados de leveduras realizado diretamente do DNA extraído podem ser visualizados na Figura 10. Analisando a foto do gel de agarose, notamos que o perfil dos fragmentos (tamanho e intensidade) dos três isolados em relação aos três genes amplificados é muito semelhante, não havendo distinção visual entre os mesmos.



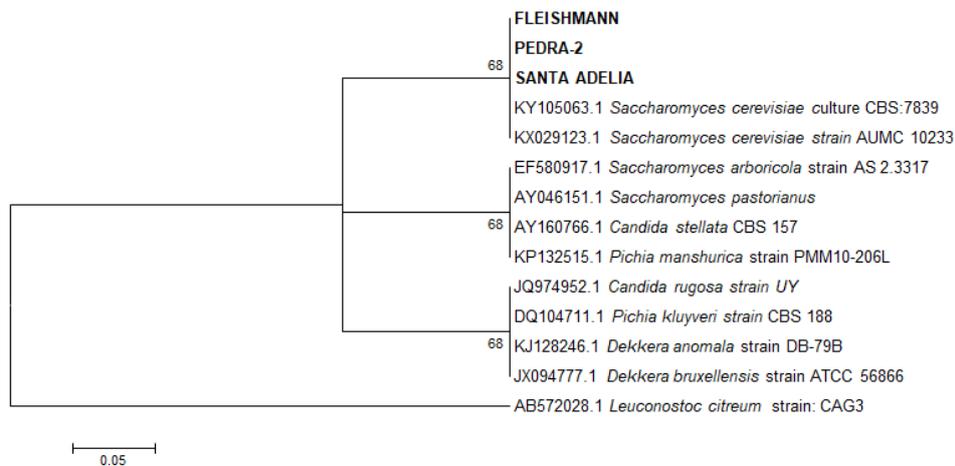
**Figura 10.** Gel de agarose (2%) com brometo de etídio (0.5 µg/mL) e a visualização por luz U.V dos fragmentos amplificados e purificados dos genes ITS, 28S e 28S (Lev) utilizando o marcador molecular Low Mass.

\*M: marcador molecular low mass; P: referente ao isolado Pedra-2; S: referente ao isolado Santa-Adelia; F: referente ao isolado Fleishmann

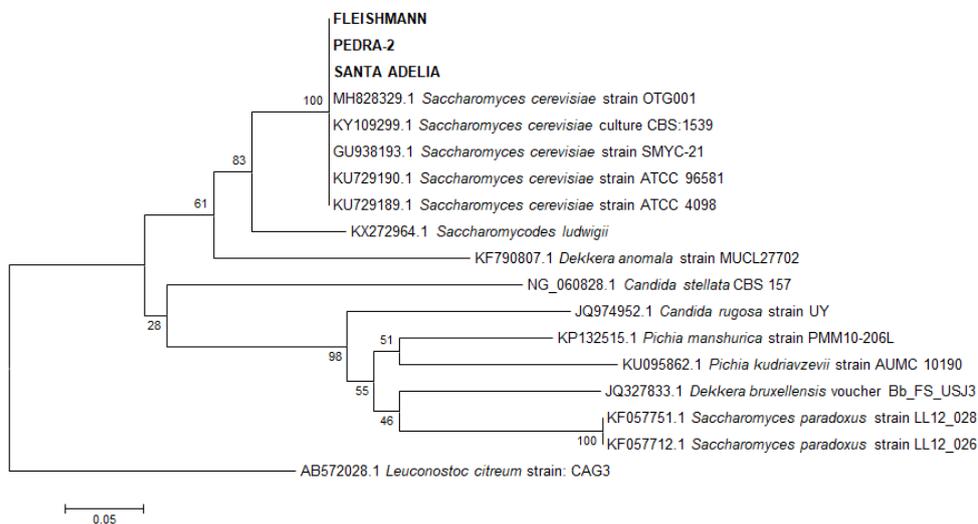
### 6.1.6. Árvore Filogenética dos isolados de *Saccharomyces*

A Figura 11 (A – C) mostra as árvores filogenéticas construídas usando sequências dos genes parciais da região ITS e 28S-nr LSU respectivamente, dos três isolados de leveduras selecionadas do processo de fermentação etanólica juntamente com cepas de *Saccharomyces* de referência recuperadas do banco de dados de nucleotídeos (<http://www.ncbi.nlm.nih.gov/blast>). A escolha de acrescentar sequências do gênero *Candida*, *Pichia*, *Dekkera* e *Saccharomyces* justifica-se por ser o gênero de leveduras contaminantes mais comumente encontrados em processos fermentativos (FLEET, 2008). As topologias das árvores filogenéticas, geradas pelo método Neighbor-Joining, foram confirmadas pela topologia de árvore gerada pelo método da máxima verossimilhança.

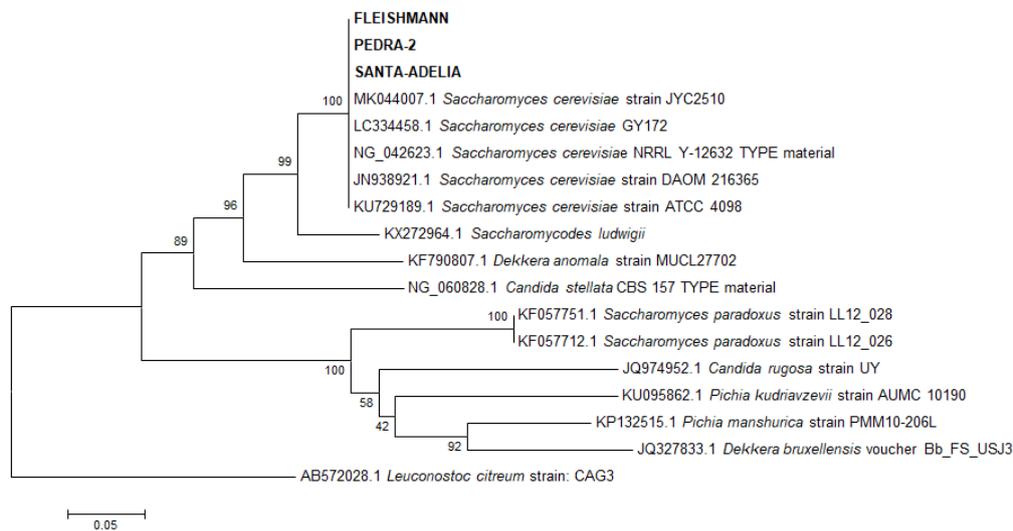
A)



B)



C)



**Figura 11.** A) Árvore filogenética baseada nas sequências do gene ITS construída usando o método neighbor-joining. *Leuconostoc citreum* CAG3 foi selecionado como grupo de fora (*out group*) e 5% substituições por sítio. B) Árvore filogenética utilizando a sequência parcial do gene 28S-nr LSU com os primers U1/U2, construída usando o método neighbor-joining. *Leuconostoc citreum* CAG3 foi selecionado como grupo de fora (*out group*) e 5% substituições por sítio. C) Árvore filogenética utilizando a sequência parcial do gene 28S-nr LSU com os primers D1/D2. *Leuconostoc citreum* CAG3 foi selecionado como grupo de fora (*out group*) e 5% substituições por sítio

\*Os valores de bootstrap maiores que 60% estão mostrados nos internós. Nomes em negrito e com letras maiúsculas pertencem às leveduras isoladas deste estudo. Siglas sem negrito pertencem ao código da sequência pertencente ao banco de dados de nucleotídeos.

Analisando as árvores filogenéticas (Figura 11 A-C) notamos que a topologia foi semelhante para ambos os genes, ITS e 28S, independente da região amplificada, ficando os três isolados deste estudo no mesmo nó juntamente com apenas cepas de *S. cerevisiae*. Resultados encontrados por Segura e colaboradores (2010) sugeriram que normalmente as regiões ribossomais 5.8S, 18S e 28S não são discriminatórias o suficiente quando comparava indivíduos de espécies intimamente relacionadas, no entanto, as regiões ITS possuíam maior variabilidade na sequência, permitindo a identificação interespecífica (GARDES et al., 1991; HENRY et al., 2000; SOUZA, 2010). Diferentemente, neste estudo notamos que tanto a região ITS quanto o gene parcial 28S foram discriminatórias o suficiente para classificar os três isolados como *S. cerevisiae*, entretanto, não foi possível distinguir ao nível de estirpe.

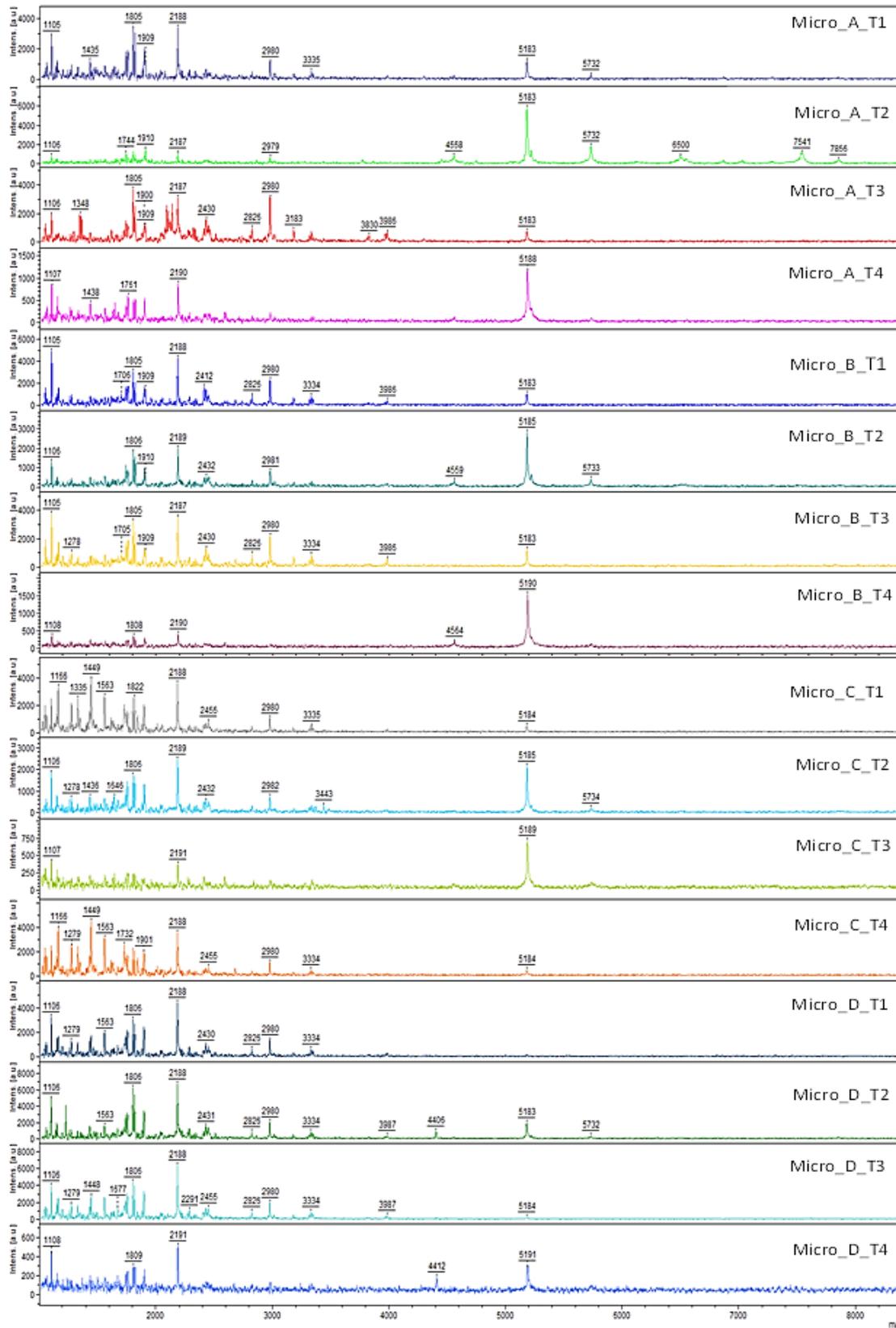
Estudos realizados por Schoch e colaboradores (2002) testaram seis regiões do DNA visando encontrar potenciais marcadores que fossem universais para o reino *Fungi* (DNA *barcode*): três subunidades cistrônicas do RNA ribossômico nuclear (18S, 5.8S e 28S) foram comparadas com três genes representativos que codificavam proteínas (subunidade maior da RNA polimerase II, segunda maior subunidade da RNA polimerase II e proteína de manutenção do minicromossomo) e concluíram que embora regiões codificadoras de proteínas frequentemente tenham uma maior porcentagem de identificação correta quando comparadas com marcadores ribossômicos, alguns entraves como pouca amplificação no PCR, dificultando o sequenciamento, fizeram com que a região ITS (ITS 1 + ITS 2 + 5.8S) fosse bem-sucedida para uma gama mais ampla de fungos, com uma variabilidade na sequência mais claramente definida para classificações inter e intraespecíficas. Além disso, em comparação com as subunidades

nucleares, a subunidade maior (28S) teve um poder discriminatório ligeiramente menor que a região ITS, sendo muitas vezes utilizada sozinha para a identificação de leveduras ou concomitantemente com a região ITS (Schoch et al., 2002); já a subunidade menor (18S), teve um poder discriminatório significativamente inferior ao ITS e 28S. Estes resultados sustentam os resultados encontrados pelas árvores filogenéticas baseadas na região ITS e 28S, que classificaram os três isolados como *S. cerevisiae*.

## **6.2. Preparo de amostras para identificação dos espectros referentes aos diferentes microrganismos utilizando o sistema MALDI-TOF Bruker Microflex**

### **6.2.1. Otimização do preparo das amostras bacterianas para aquisição por MALDI-TOF MS**

Para a escolha do modo de aquisição do perfil espectral dos isolados bacterianos, selecionamos quatro isolados de maneira aleatória, sem identificação prévia (denominados de Microrganismo A, B, C e D), os quais foram submetidos a quatro diferentes metodologias de extração proteica, diferentes modos de aplicação da amostra na placa do MALDI e diferentes concentrações das matrizes. Os resultados dos espectros de cada isolado utilizando diferentes tratamentos (denominados de T1, T2, T3, T4) podem ser visualizados na Figura 12. Foram avaliados cada espectro individualmente levando em consideração a qualidade do espectro obtido, o grau de ruído, número de picos e a reprodutibilidade em relação aos diferentes microrganismos, pois o método de extração escolhido teria que produzir espectros com qualidade suficiente para uma boa identificação, independentemente da espécie do isolado. Posteriormente, os dados brutos foram contrastados com os espectros do banco de dados MALDI-Biotyper para verificar qual tratamento obtinha a maior nota ( $\log(\text{score})$ ) e, conseqüentemente, o melhor método de identificação. Os valores das notas obtidas pelo Biotyper dos microrganismos A, B, C e D, separadamente, submetidos aos diferentes tratamentos (T1, T2, T3 e T4) estão no Gráfico 1.

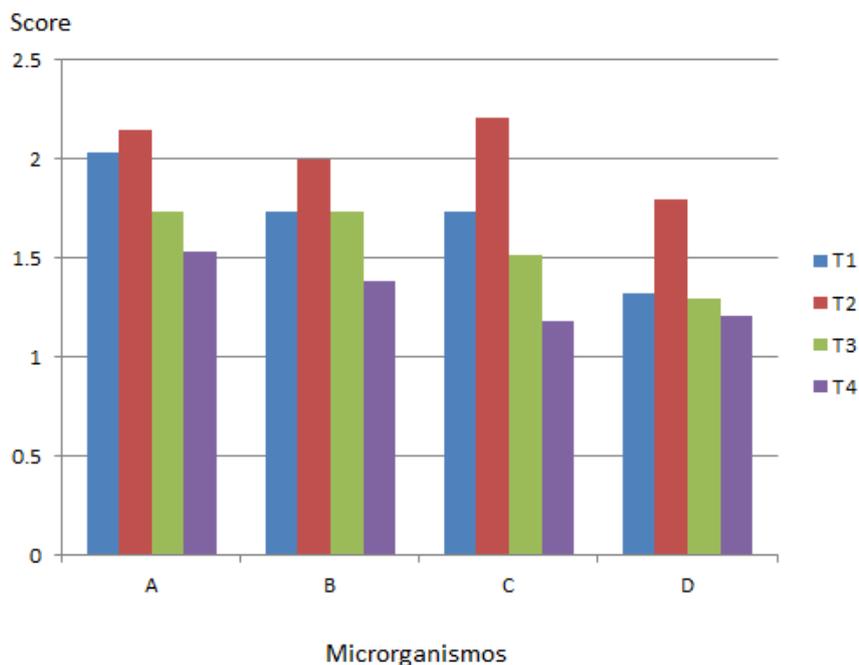


**Figura 12.** Comparação do perfil proteico espectral dos microrganismos A, B, C e D submetidos a diferentes metodologias de extração.

\*Ordenada = intensidade relativa; abcissa = massa / carga ( $m/z$ ); T1 = tratamento número um; T2 = tratamento número dois; T3 = tratamento número três, T4 = tratamento número quatro.

Analisando a Figura 12, notamos que o tratamento dois (T2), método baseado em Posteraro e colaboradores, (2012) foi o método em que mais picos foram identificados, com menor ruído e com maior qualidade, principalmente quando analisamos os microrganismos C e D, talvez por ser o único método que possua uma etapa adicional de adição de etanol absoluto entre a aplicação da amostra e da matriz na placa, permitindo, possivelmente, uma maior secagem e limpeza da amostra, possibilitando uma melhor cristalização da matriz e, consequentemente, uma ionização mais branda e eficiente da amostra. Além disso, foi o método de extração com menor número de passos, sem centrifugações, permitindo o desenvolvimento de uma metodologia mais robusta e simplificada. Já o tratamento quatro (T4), método baseado em Alatoon e colaboradores, (2011) teve a pior performance em todos os quatro isolados, com o perfil espectral com menor qualidade. Provavelmente, por ser o método com o maior número de lavagens e centrifugações e, além disso, ser o único a utilizar a matriz em solução saturada com apenas 30% de acetetonitrila e 0.1% de ácido trifluoroacético, o que pode ter prejudicado a ionização, já que o ácido trifluoroacético auxilia na ionização de moléculas com alta massa molecular, como as proteínas, que são naturalmente mais difíceis de serem ionizadas. Já nos outros tratamentos foram utilizadas quantidades maiores dos dois reagentes, 2,5% de ácido trifluoroacético e 50% de acetoneitrila, o que pode ter melhorado a ionização.

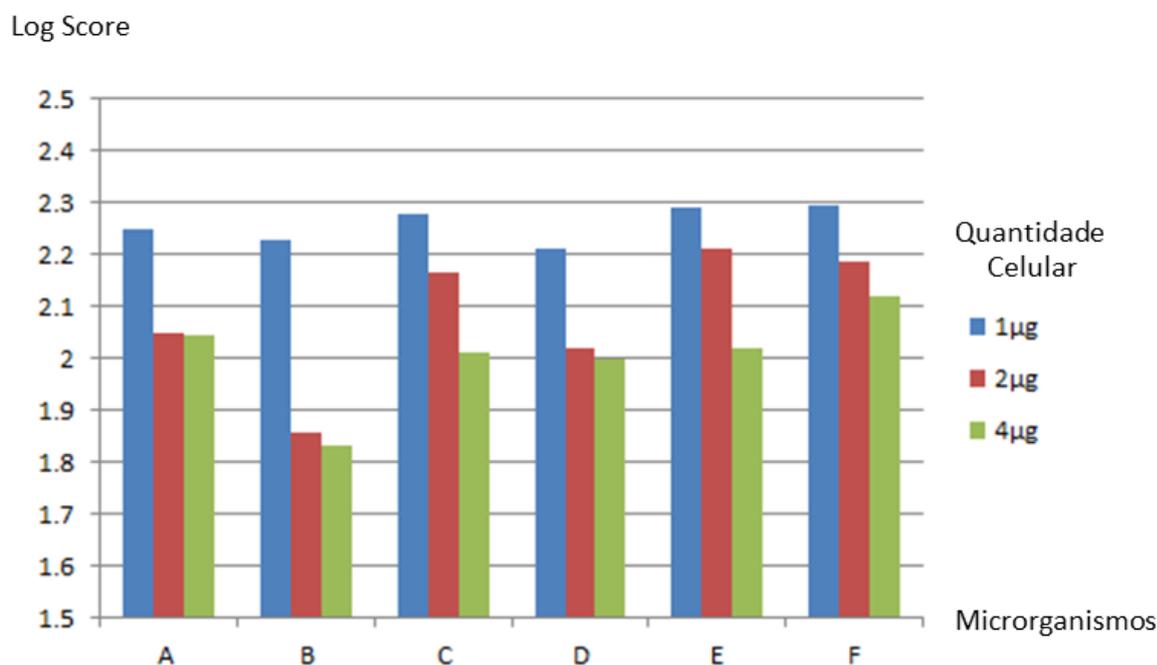
Relacionando os espectros de massa (Figura 12) com suas respectivas notas pelo banco de dados Biotyper (Gráfico 1), indubitavelmente nota-se que a qualidade do espectro influenciou diretamente no poder classificatório do banco de dados, sendo que o T2, com maior qualidade espectral, teve consequentemente, as maiores notas de simetria entre os espectros dos quatro microrganismos e os espectros de referência; já o T4, com a pior qualidade espectral, também teve as piores notas pelo banco de dados Biotyper.



**Gráfico 1:** Correlação entre o perfil espectral obtido por MALDI-TOF MS dos microrganismos A, B, C e D submetidos a diferentes tipos de preparo e aplicação de amostra (T1, T2, T3 e T4) e as notas correspondentes das identificações realizadas pelo banco de dados Biotyper.

\*Ordenada: valor de log score do banco de dados Biotyper em relação a simetria espectral entre os espectros dos microrganismos A, B, C e D e os espectros do banco de dados; Abcissa: identificação dos microrganismos; Tn: refere-se aos diferentes preparos e aplicações de amostra para aquisição dos espectros de massa.

Após a escolha do melhor método para aquisição do espectro foram realizados testes para verificar a quantidade de células ideais (1 µg, 2 µg ou 4 µg) para obter um espectro com qualidade, permitindo a identificação pelo software biotyper. Para os testes foram selecionados os quatro microrganismos utilizados para os testes anteriores (A, B, C e D) e foram adicionados outros dois microrganismos (E e F). Os dados brutos adquiridos por MALDI-TOF MS foram adicionados ao Biotyper para a identificação espectral e de acordo com o grau da simetria dos espectros dos isolados com os espectros de referências do banco, remetia um score de 0 (mínimo score) a 3 (máximo score). Os valores das notas obtidas por cada um dos microrganismos em diferentes concentrações celulares, 1 µg, 2 µg e 4 µg, podem ser visualizados no Gráfico 2.



**Gráfico 2:** Correlação entre o valor de score obtido pelo banco Biotyper e a quantidade celular de cada microrganismo na obtenção do perfil espectral por MALDI-TOF MS.

\*Ordenada: valor de log score do banco de dados Biotyper em relação a simetria espectral entre os espectros dos microrganismos A, B, C, D, E e F e os espectros do banco de dados; Abcissa: identificação dos microrganismos; Quantidade celular: refere-se a quantidade de células inicial utilizada para realizar a extração proteica.

Analisando o gráfico 2 notamos que o valor do score obtido pelo Biotyper foi maior quando usamos uma quantidade menor de células (1 µg) para todos os microrganismos. Além disso, notavelmente houve um decréscimo da simetria espectral em relação a identificação utilizando o banco com o aumento da concentração celular. Estudos realizados por Moothoo-Padayachie e colaboradores (2002), tentando otimizar o preparo de amostra e aquisição espectral de isolados de *S. cerevisiae* também testou o efeito da quantidade de células no poder de identificação do Biotyper e notou que com o aumento celular houve um decréscimo do score remetido pelo banco de dados, especificamente, quando a quantidade de células foi acima de  $10^8$ . Por se tratar de um equipamento de alta sensibilidade e a qualidade do espectro ser altamente dependente da polimerização da matriz na amostra, uma grande quantidade celular pode saturar a matriz, não havendo uma polimerização adequada, afetando negativamente a aquisição espectral quando comparado com espectros de banco de dados. Entretanto, apesar da diminuição do score, não houve perda substancial na identificação, sendo que todas as concentrações celulares utilizadas remeteram a classificação da mesma espécie de microrganismo. Desta forma, correlacionando os resultados obtidos do gráfico 2,

concluimos que quando a extração de proteínas foi feita a partir de 1 µg de células, obteve-se o melhor resultado em relação a identificação espectral.

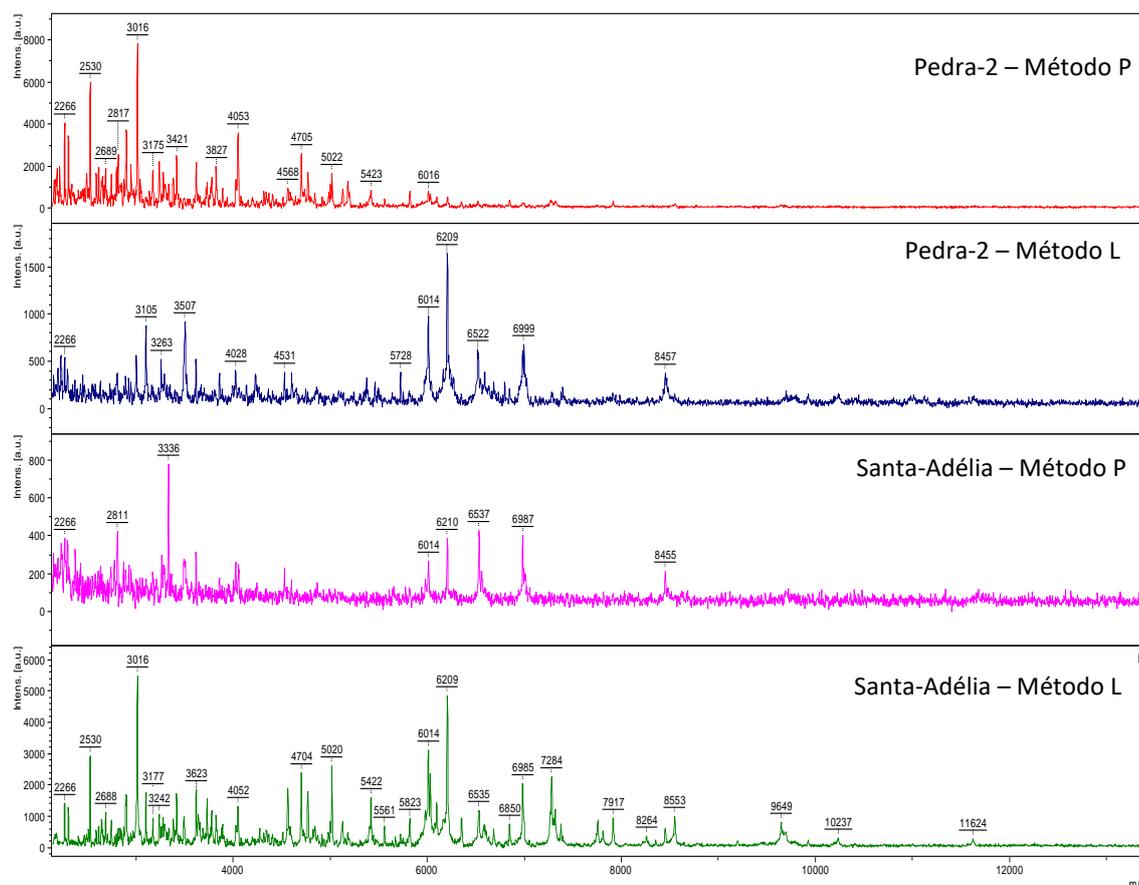
### 6.2.2. Preparo das amostras de leveduras para aquisição por MALDI-TOF MS

Os resultados da aquisição dos espectros dos isolados de leveduras Pedra-2 e Santa-adélia submetidos a dois métodos de extração proteica, baseado em Posteraro e colaboradores (2012), denominado como Método-P e, em Lauterbach e colaboradores (2017), denominado como Método-L, podem ser visualizados a Figura 13 (A e B).

Analisando o perfil espectral de ambos isolados, notamos claramente que o método-P utilizado para bactérias não obteve o mesmo resultado satisfatório para leveduras e que o método-L foi o que mais picos foram identificados, com menos ruído e com maior qualidade, principalmente para o isolado Santa-Adélia. Essa diferença, provavelmente é justificada pelas diferenças moleculares e estruturais entre microrganismos procariotos e eucariotos.

Estruturalmente, as bactérias possuem paredes celulares compostas por peptidoglicano (90% da composição em bactérias gram-positivas e 10% em bactérias gram-negativas), um polissacarídeo formado por cadeias alternadas de N-acetilglucosamina e ácido N-acetilmurâmico que são interligadas por meio de ligações peptídicas cruzadas. Já as paredes das leveduras são compostas principalmente por polissacarídeos de glicose ( $\beta$ -1,3- e  $\beta$ -1,6-glucanas), manose (ligados a proteínas na parte externa da parede) e N-acetilglicosamia (em pequena quantidade). Outra diferença está na estruturação celular, células bacterianas não possuem núcleo compartimentalizado, possuindo o DNA condensado (nucleóide) sem membrana nuclear, não havendo separação do material genético com o citoplasma, desta forma todo o DNA, ribossomos e proteínas ficam expostos no citoplasma. Já as células de leveduras, por serem eucarióticas, possuem núcleo com membrana, separado do citoplasma, DNA condensado na forma de cromatina e compartimentalização celular (Antonini, 2012). Molecularmente, pela ausência de compartimentalização, nas bactérias o processo de transcrição e tradução ocorrem simultaneamente, ou seja, as proteínas são produzidas juntamente com a leitura do RNA mensageiro. Já em leveduras, enquanto a transcrição ocorre no núcleo, a tradução ocorre no citoplasma, havendo o processamento do RNA mensageiro antes de ser transformado em cadeia polipeptídica, além de possuir membranas com seu próprio DNA, como mitocôndrias e várias proteínas dentro de outros compartimentos celulares. Desta forma, considerando todas as peculiaridades citadas de microrganismos eucariotos, o método de extração proteica mais elaborado, com etapas de rompimento celular, lavagens e centrifugações e com maior concentração de ácido fórmico permitiu uma aquisição espectral mais eficiente.

A)



B)

Microrganismo	Método	Log score	Classificação Biotyper
Pedra-2	P	1.97	Provável <i>Saccharomyces</i>
Pedra-2	L	2.30	<i>S. cerevisiae</i>
Santa-Adélia	P	2.12	<i>Saccharomyces</i>
Santa-Adélia	L	2.33	<i>S. cerevisiae</i>

**Figura 13.** A) Comparação do perfil proteico espectral dos isolados de leveduras Pedra-2 e Santa-Adélia submetidos a diferentes metodologias de extração. B) O valor de log score obtido pelo MALDI-Biotyper de cada isolado de acordo com o método de extração utilizado e suas respectivas classificações.

\*Ordenada = intensidade relativa; abscissa = massa / carga ( $m/z$ ); T1 = tratamento número um; T2 = tratamento número dois; T3 = tratamento número três; T4 = tratamento número quatro.

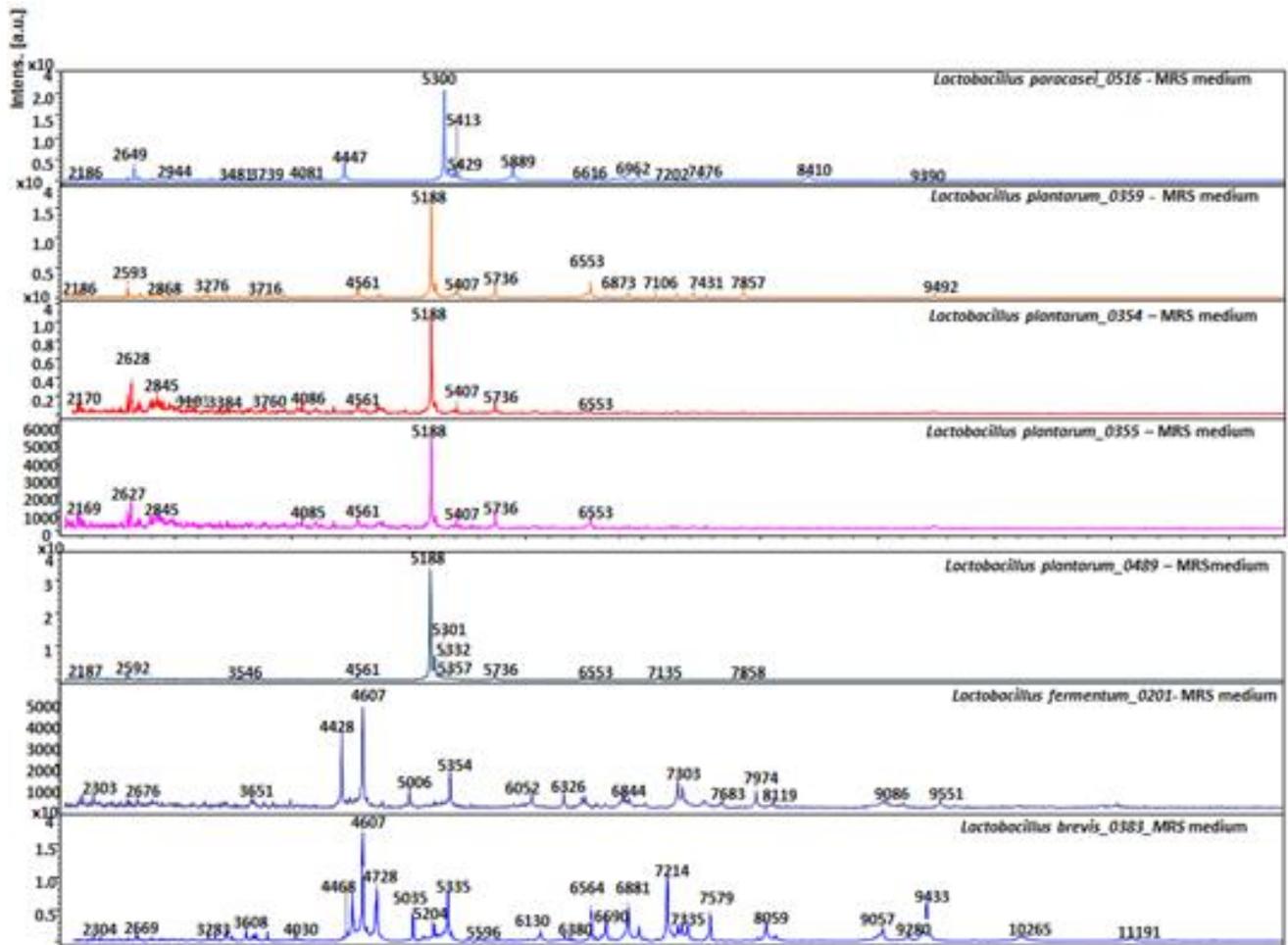
### 6.3. Aquisição dos espectros de *Lactobacillus* e análise dos dados obtidos por MALDI-

#### TOF MS

#### 6.3.1. Perfil proteico espectral de isolados de *Lactobacillus* comerciais em meio MRS para validação da técnica MALDI-TOF para identificação microbiológica

Para verificar se a técnica era eficaz para discriminar isolados de *Lactobacillus*, alguns isolados disponíveis no banco de dados de Lavras (CCMA), previamente identificados, foram submetidos a extração proteica, preparo e

aplicação da amostra baseados no método de Posteraro e colaboradores (2012), o qual teve melhor performance para bactérias e os resultados dos perfis espectrais podem ser visualizados na Figura 14.



**Figura 14.** Perfil spectral dos isolados de *Lactobacillus* pertencentes a coleção microbiológica de Lavras (CCMA) obtido por MALDI-TOF a partir do crescimento microbiano em meio MRS.

\*Ordenada: intensidade relativa; abscissa: massa / carga ( $m/z$ ); Nomes no canto superior esquerdo: classificação dada pelo banco de dados de Lavras.

Notavelmente os espectros de massas forneceram perfis espectrais característicos para cada espécie analisada, sendo semelhantes entre isolados da mesma espécie, evidenciando um padrão espectral. Entretanto, quando utilizamos o banco de dados Biotyper como referência para a identificação por similaridade espectral, dos 7 isolados testados, apenas os isolados 0516 e 0359 possuíam um score alto o suficiente que permitia confiabilidade na discriminação a nível de espécie, todos os outros isolados tiveram notas que permitiram apenas confiabilidade ao nível de gênero, *Lactobacillus*. Esses resultados mostram que as impressões digitais ao nível de gênero são mais consistentes, permitindo uma identificação confiável por similaridade espectral de maneira mais robusta, mesmo quando os espectros do banco de dados provieram de microrganismos clínicos. Entretanto, para a identificação com alta confiabilidade a nível de espécie, é necessária uma sobreposição espectral mais similar entre o número de  $m/z$  e as intensidades dos picos, discernindo entre isolados de diferentes ambientes. Contudo, é necessário salientar que apesar do score não ter sido alto o suficiente para a identificação precisa ao nível de espécie, estipulado pelo Biotyper, todas as espécies sugeridas com o score para identificação ao nível de gênero foram correspondentes. Essa contestação só foi possível devido a identificação gênica prévia dos microrganismos, se tratassem de microrganismos

desconhecidos da área agrícola, dificilmente teríamos uma identificação confiável ao nível de espécie apenas com os microrganismos já depositados no banco de dados Biotyper.

### 6.3.2. Perfil espectral de isolados de *Lactobacillus* comerciais em diferentes meios de cultura

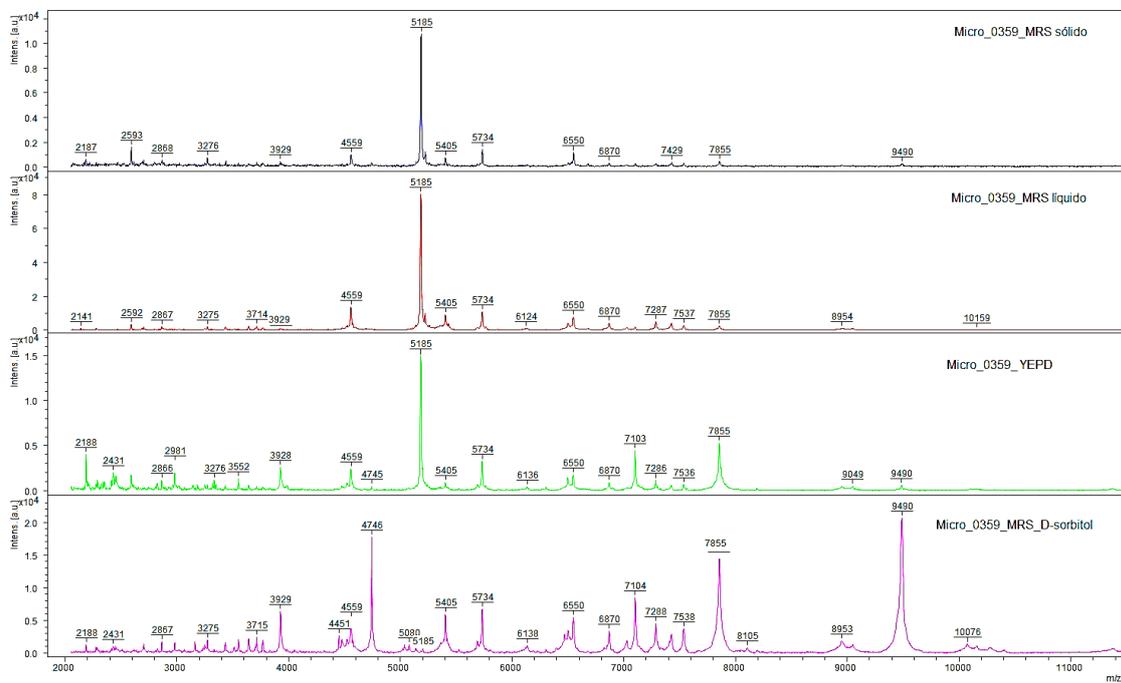
Os perfis espectrais dos três microrganismos comerciais adquiridos em diferentes meios de cultura podem ser visualizados na Figura 15 (A - C) com suas respectivas notas e classificações pelo banco de dados Biotyper (Figura 15 - D).

Analisando os perfis espectrais dos microrganismos adquiridos em diferentes meios, notamos que houve regiões variáveis de acordo com o meio de cultura utilizado, principalmente quando a fonte de açúcar foi trocada pelo D-sorbitol. Contudo, grande parte dos  $m/z$  dos picos mantiveram-se os mesmos, sendo característicos do microrganismo, com variação apenas na intensidade relativa. Estudos fisiológicos e proteômicos realizados por Long e colaboradores (2002) indicaram que diferentes nutrientes no meio de cultura poderiam afetar o padrão de proteases em bactérias da espécie *Bacillus subtilis* e *Serratia marcescens*. Arnold e colaboradores (1999) mostrou que havia variações nos perfis espectrais de *Escherichia coli* quando estes eram adquiridos ao longo da curva de crescimento microbiano. Além disso, de acordo com estudos com proteômica utilizando eletroforese bidimensional, Choe, Chen e Lee (1999) constataram que *Bacillus subtilis* e *Escherichia coli* quando crescidos em meio mínimo ocorriam mudanças na expressão proteica e na composição do envelope celular em comparação com o meio complexo e produção de diferentes polímeros da parede celular. Particularmente, os níveis de galactose e galactosamina variaram em estirpes de *B. subtilis* cultivadas em meio rico nutritivo versus meio mínimo semidefinido com fosfato, demonstrando a variabilidade na composição celular (FOX et al., 2000). Adicionalmente, outros estudos relataram numerosos genes cuja expressão foi alterada pelo crescimento em meio com uma fonte mínima de sais ou carbono versus um meio rico complexo (PETERSOHN et al., 2001).

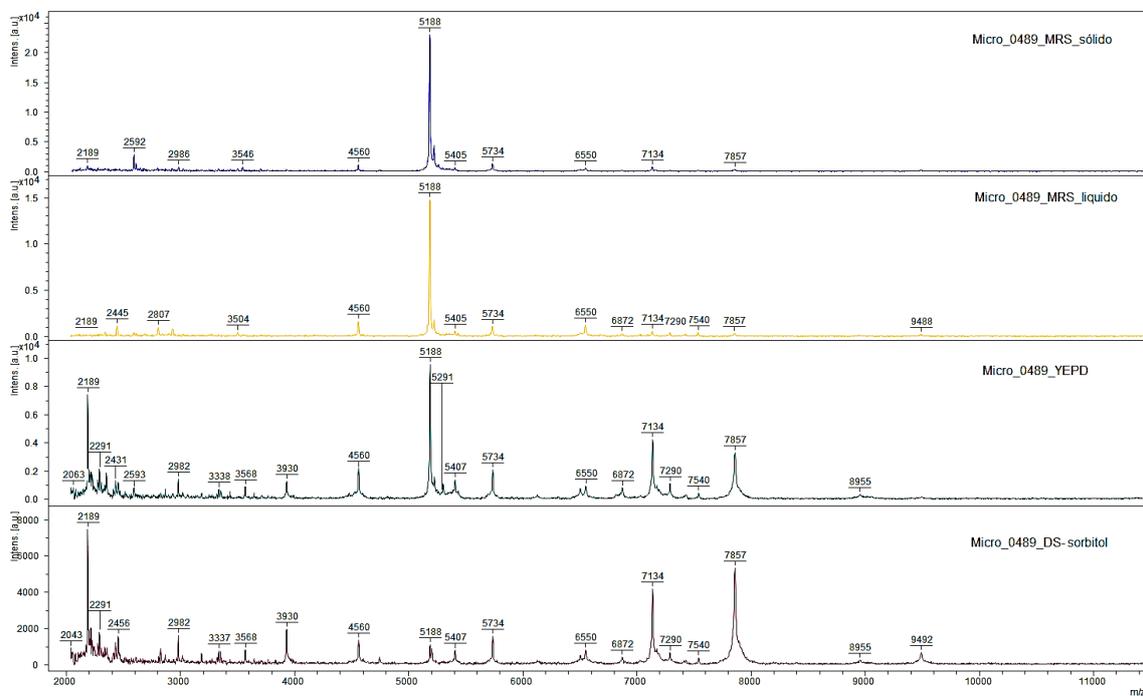
Em contrapartida, Bernardo e colaboradores (2002) ao crescer bactérias do gênero *Staphylococcus* em dois diferentes meios de cultura mostrou que não houve alteração significativa nos perfis de MALDI-TOF MS. E, posteriormente estudos realizados por Valentine e colaboradores (2005) com diferentes bactérias constataram que enquanto algumas proteínas diferentes são produzidas por bactérias vegetativas quando são cultivadas em diferentes meios de crescimento, a identificação positiva com MALDI-TOF MS ainda foi possível e, concluíram que havia um conjunto de proteínas nucleares que permaneciam constantes no espectro mesmo quando as bactérias eram crescidas em quatro diferentes meios de cultura, desde o meio mínimo até o mais complexo, e que as análises das proteínas intactas por MALDI-TOF MS mostrou-se ser uma técnica de identificação consistente.

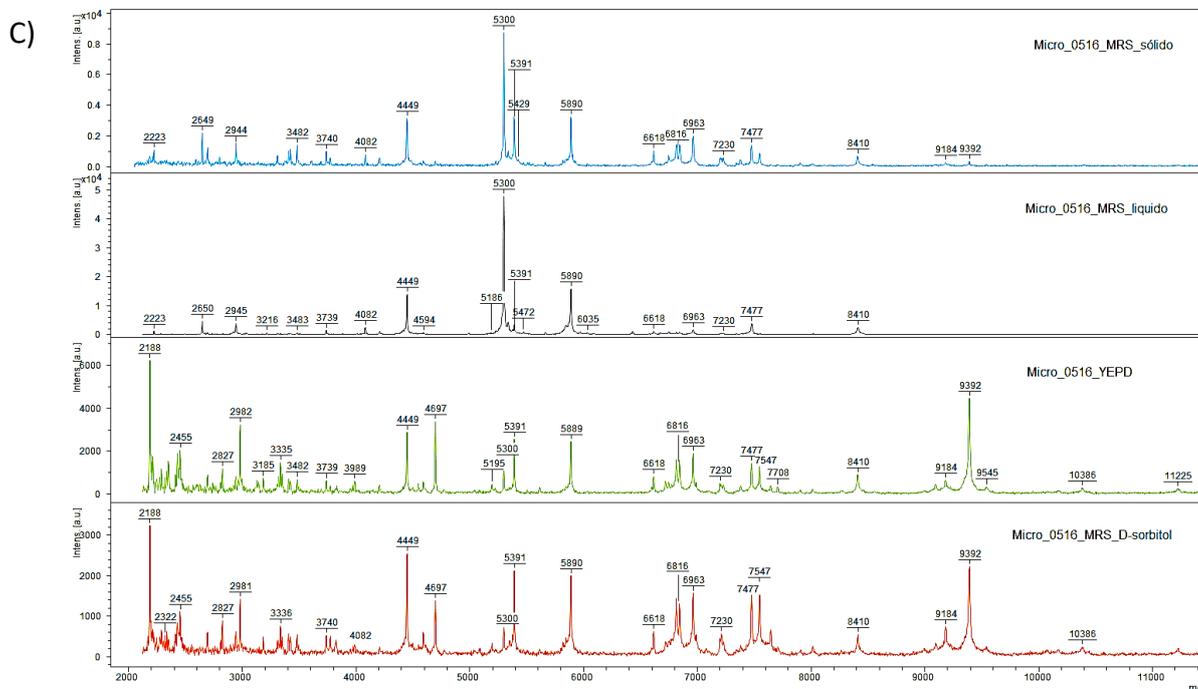
Os resultados e conclusões de Valentine e colaboradores (2005) corroboram com os resultados encontrados neste estudo, pois por mais que houvessem variações em regiões do espectro (Figura 15 A- C), quando os espectros foram comparados com o banco de dados espectral do Biotyper (Figura 15 -D) remeteram a identificação da mesma espécie, não havendo perda substancial na identificação. Além disso, já foi constatado que a maioria das proteínas ionizadas cujo sinal é detectado por MALDI-TOF, é de origem ribossomal sendo uma técnica eficiente para biomarcadores.

A)



B)





D)

Meios de Cultura	CCMA-0359		CCMA-0489		CCMA0516	
	(Log) Score Biotyper	Classificação Biotyper	(Log) Score Biotyper	Classificação Biotyper	(Log) Score Biotyper	Classificação Biotyper
MRS sólido	2.427	<i>L. plantarum</i> 0489	2.358	<i>L. plantarum</i> 0359	2.386	<i>L. paracasei</i> DSM-20312
MRS líquido	2.430	<i>L. plantarum</i> DSM-2601	2.342	<i>L. plantarum</i> 0359	2.338	<i>L. paracasei</i> DSM-20312
YEPD	2.342	<i>L. plantarum</i> DSM-2601	2.346	<i>L. plantarum</i> 0359	2.331	<i>L. paracasei</i> DSM-20244
MRS- D-sorbitol	2.251	<i>L. plantarum</i> DSM-1055	2.060	<i>L. plantarum</i> DSM-1055	2.330	<i>L. paracasei</i> DSM-20244

**Figura 15.** Perfil espectral dos microrganismos 0359, 0439, 0516 adquiridos por MALDI-TOF MS em quatro meios de cultura: meio MRS sólido, meio MRS líquido, meio YEPD e meio MRS D-sorbitol e suas respectivas notas e classificações utilizando o banco de dados Biotyper.

Proteínas ribossomais podem ser usadas como biomarcadores por possuírem alta conservação genética e expressão constitutiva, não variando de acordo com o meio (TERAMOTO et al., 2007). Esta alta conservação genética está relacionada a função primordial destas proteínas, as quais auxiliam na conformação tridimensional do ribossomo para que a síntese protéica ocorra de forma eficiente. Entretanto, podem aparecer eventualmente outras proteínas que estão presentes em maior quantidade (TERAMOTO et al., 2007; ASSIS, 2011), o que justifica os espectros diferenciais da Figura 15.

Além disso, do ponto de vista molecular, organismos procarióticos possuem características que podem ser agravantes na obtenção do espectro, tais como: a) o genoma não está organizado em cromossomos, estando

prontamente disponível para a transcrição; assim, enquanto em organismos eucarióticos os genomas estão “desligados” a não ser que ativados, em procarióticos o genoma sempre está “ligado” a não ser que reprimidos por algum mecanismo de regulação; b) o genoma está organizado em operón, sendo que um promotor é responsável pela regulação de uma cascata de genes intimamente relacionados, os quais são transcritos em um único RNA mensageiro (RNAm); c) conseqüentemente, os principais mecanismos de regulação gênica são pré-transcricionais; e, d) por não possuírem compartimentalização celular, conseqüentemente a informação contida na molécula de DNA é transcrita e traduzida em proteínas quase que instantaneamente, proporcionando uma rápida resposta fisiológica e adaptação em relação ao meio.

Por fim, apesar das diferenças nas assinaturas espectrais serem observadas quando o mesmo microrganismo foi crescido em diferentes meios, ressalta-se que a identificação não foi ambígua ou seriamente prejudicada pelo banco de dados Biotyper, pois ainda persistiram íons diagnósticos relevantes que permitiram a identificação correta das cepas de *Lactobacillus*. Entretanto, a partir dos dados obtidos nesse estudo e no interesse de manter um protocolo padronizado para criação de um banco de dados, é aconselhável utilizar o mesmo meio de cultivo para a biotipagem de isolados de *Lactobacillus*.

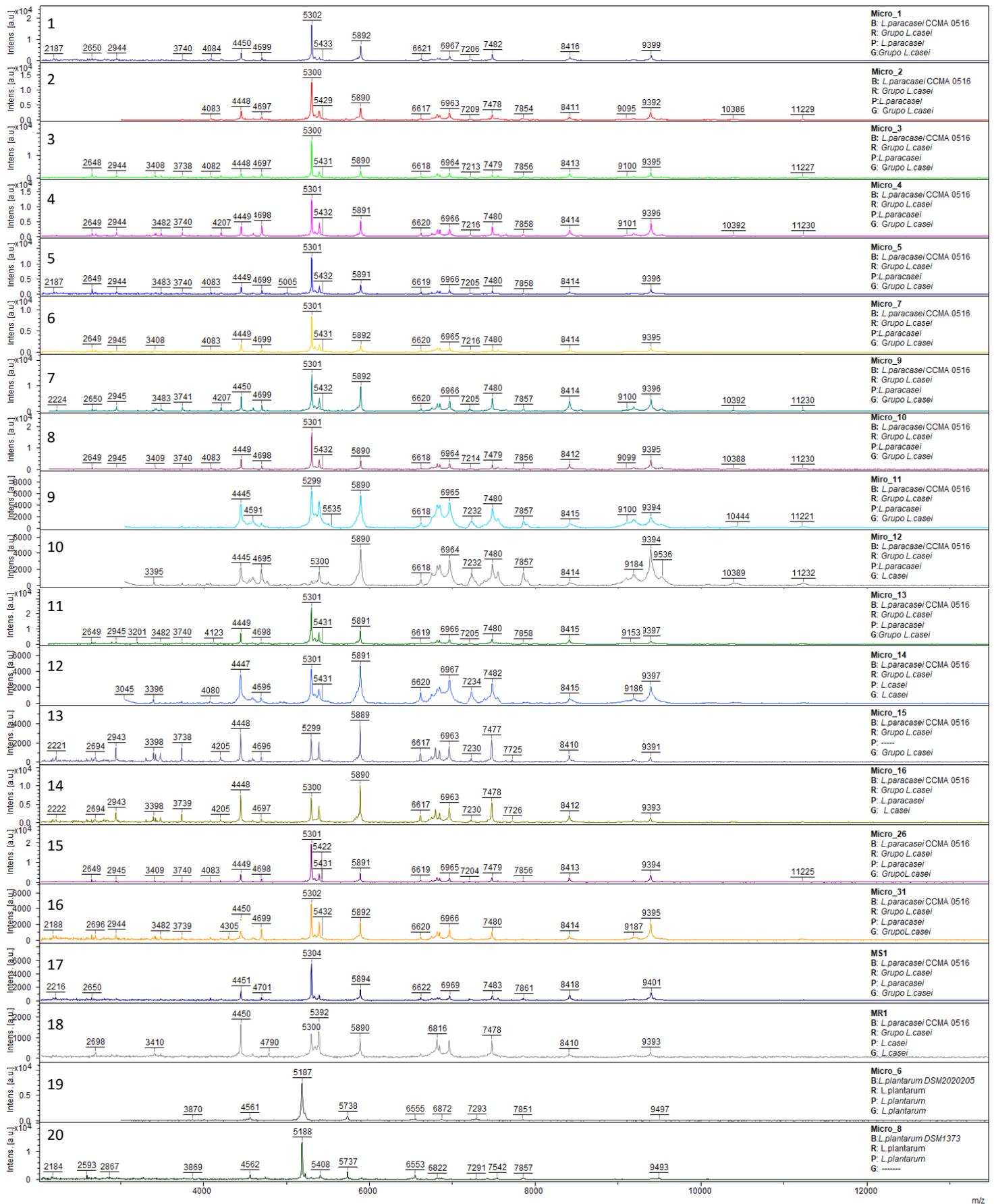
### 6.3.3. Perfil proteico espectral dos isolados de *Lactobacillus* do processo de fermentação etanólica

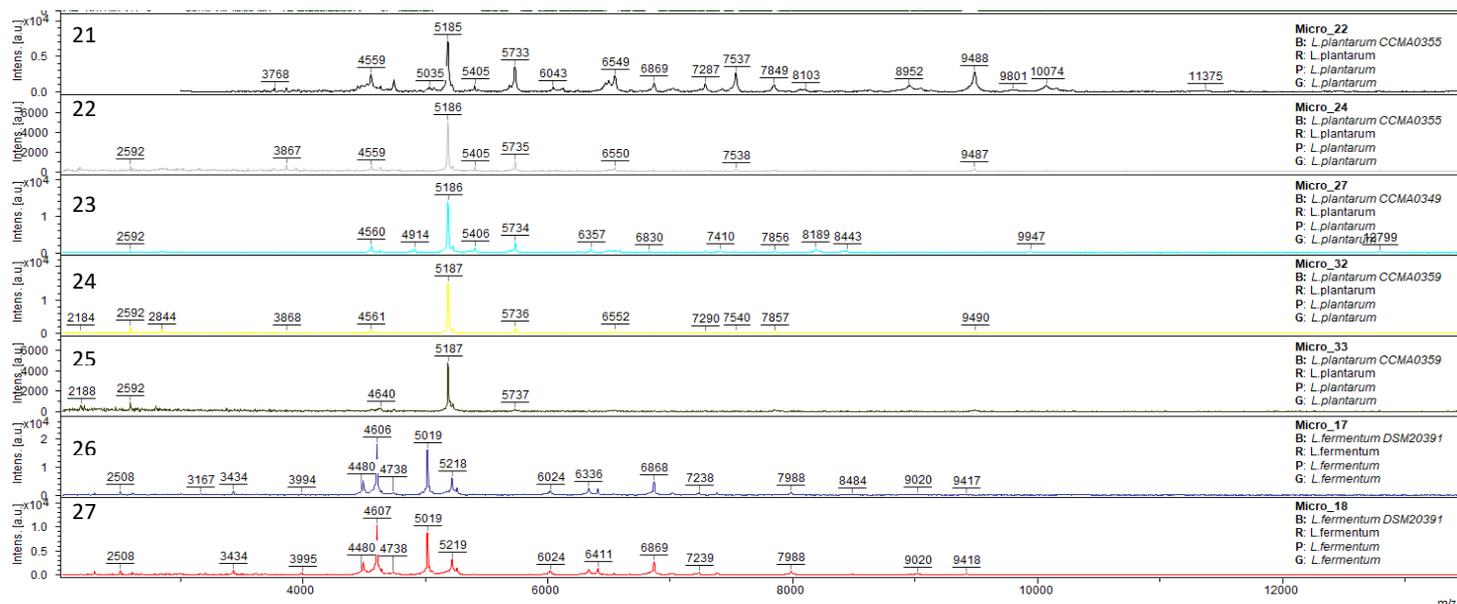
A figura 16 mostra o perfil espectral obtido por MALDI-TOF de todos os isolados do processo de produção de etanol pertencentes ao gênero *Lactobacillus*, crescidos previamente em meio seletivo MRS com ágar, com suas respectivas identificações utilizando o banco de dados espectral Biotyper, acrescidos os espectros dos sete microrganismos comerciais em um banco de dados complementar interno ao software e pelas árvores filogenéticas baseadas nos genes 16S rDNA, *pbES* e *groEL*.

Nota-se que os espectros de massas de bactérias forneceram impressões espectrais específicas que permitiram o agrupamento de espécies em relação aos valores de  $m/z$ , número e intensidade dos picos semelhantes, sendo pelo menos um pico específico com alta intensidade foi observado para cada espécie (Essas características podem ser visualizadas de forma mais detalhada na Figura 17). Levando-se em consideração a classificação obtida pelo banco de dados Biotyper (Figura 16, canto superior direito), os isolados foram agrupados em três espécies distintas: *L.paracasei*, *L.plantarum* e *L.fermentum*. Ressalta-se que a identificação espectrométrica utilizada na figura refere-se a maior nota obtida de acordo com a similaridade do espectro do isolado do estudo em relação ao espectro de referência do banco de dados. Desta forma, notavelmente os espectros dos microrganismos agrícolas adicionados como um banco de dados interno particular obtiveram as maiores correspondências, evidenciando a particularidade espectral dos microrganismos de cada área.

Primeiramente, analisando os espectros classificados como *L. paracasei* pelo Biotyper, notamos que os picos mais intensos ( $m/z$  5.301, 5891, 4448, 9394), foram também os mais característicos dessa espécie em relação aos espectros classificados como *L.plantarum* e *L.fermentum*. Entretanto, não houve qualquer distinção entre os isolados *L.paracasei* e *L.casei*, discriminados previamente nas análises filogenéticas baseadas nos genes *housekeeping*. Estudos realizados por Dusková et al. (2012) visando a identificação de diferentes espécies de *Lactobacillus* em produtos alimentícios, também observaram que as espécies identificadas molecularmente como *L.casei* foram erroneamente classificadas como *L.paracasei* pelo banco de dados Biotyper. Uma hipótese seria devido ao banco de

dados Biotyper possuir apenas 1 isolado caracterizado como *L. casei* e nenhum isolado dessa espécie no banco de dados interno desenvolvido neste estudo.





**Figura 16.** Perfil spectral dos isolados de *Lactobacillus* do processo de produção de etanol, crescidos em meio seletivo MRS, por MALDI-TOF e suas respectivas identificações moleculares e espectrométricas.

\*Ordenada: intensidade relativa; abcissa: massa / carga ( $m/z$ ); 1-29: ordem numérica dos espectros; sigla escrito em negrito no canto superior direito: identificação do microrganismo deste estudo; B: identificação obtida pelo banco de dados espectral Biotyper; R: identificação obtida pela árvore filogenética baseada no gene 16S rDNA; P: identificação obtida pela árvore filogenética baseada no gene *pheS*; G: identificação obtida pela árvore filogenética baseada no gene *groEL*; L.: abreviação de *Lactobacillus*.

. Entretanto, há 16 isolados caracterizados como *L. paracasei* (15 pelo banco Biotyper e 1 pelo banco interno deste trabalho), havendo, portanto, uma grande variabilidade espectral dos *L. paracasei* permitindo a sobreposição espectral dos isolados deste estudo com diferentes isolados do banco de dados, o que já não é possível havendo apenas 1 isolado, pois a identificação por similaridade espectral fica restrita a apenas um espectro de referência, principalmente tratando-se de duas espécies muito próximas que possivelmente compartilham os mesmos picos característicos (por exemplo  $m/z$  4450, 5300, 5892) além de não ter uma descrição taxonômica muito clara dessa espécie (DUSKOVÁ et al., 2012). Contudo, curiosamente, podemos notar que todos os isolados classificados como *L. casei* pelo gene *groEL* (Micro\_12, Micro\_14, Micro\_15, Micro\_16 e MR1), representados pelos espectros de massa de número 10, 12, 13, 15 e 18, respectivamente (Figura 16, canto superior esquerdo) apresentaram uma característica singular no perfil espectral: o pico mais intenso foi o de  $m/z$  5891, enquanto todos os outros isolados classificados pelo gene *groEL* como *L. paracasei*, tiveram como pico mais intenso o de  $m/z$  5.300. Isso mostra que houve uma correlação negativa entre os dois picos, ou seja, sempre que um aumentava, o outro diminuía, mostrando que possivelmente as duas espécies compartilham os mesmos picos característicos, entretanto de intensidades relativas diferentes. Contudo, apesar da mudança sutil do perfil espectral e a classificação filogenética desses isolados como *L. casei* utilizando o gene *groEL*, não houve unanimidade para a classificação destes isolados quando utilizado os outros dois genes, *pheS* e 16S rDNA que reforçasse essa possível evidência, podendo apenas ser uma característica de diferentes estirpes de *L. paracasei*. Desta forma, evidenciamos que tanto geneticamente, quanto espectrometricamente, os isolados pertencentes ao grupo *L. casei* são muito difíceis de serem discriminados.

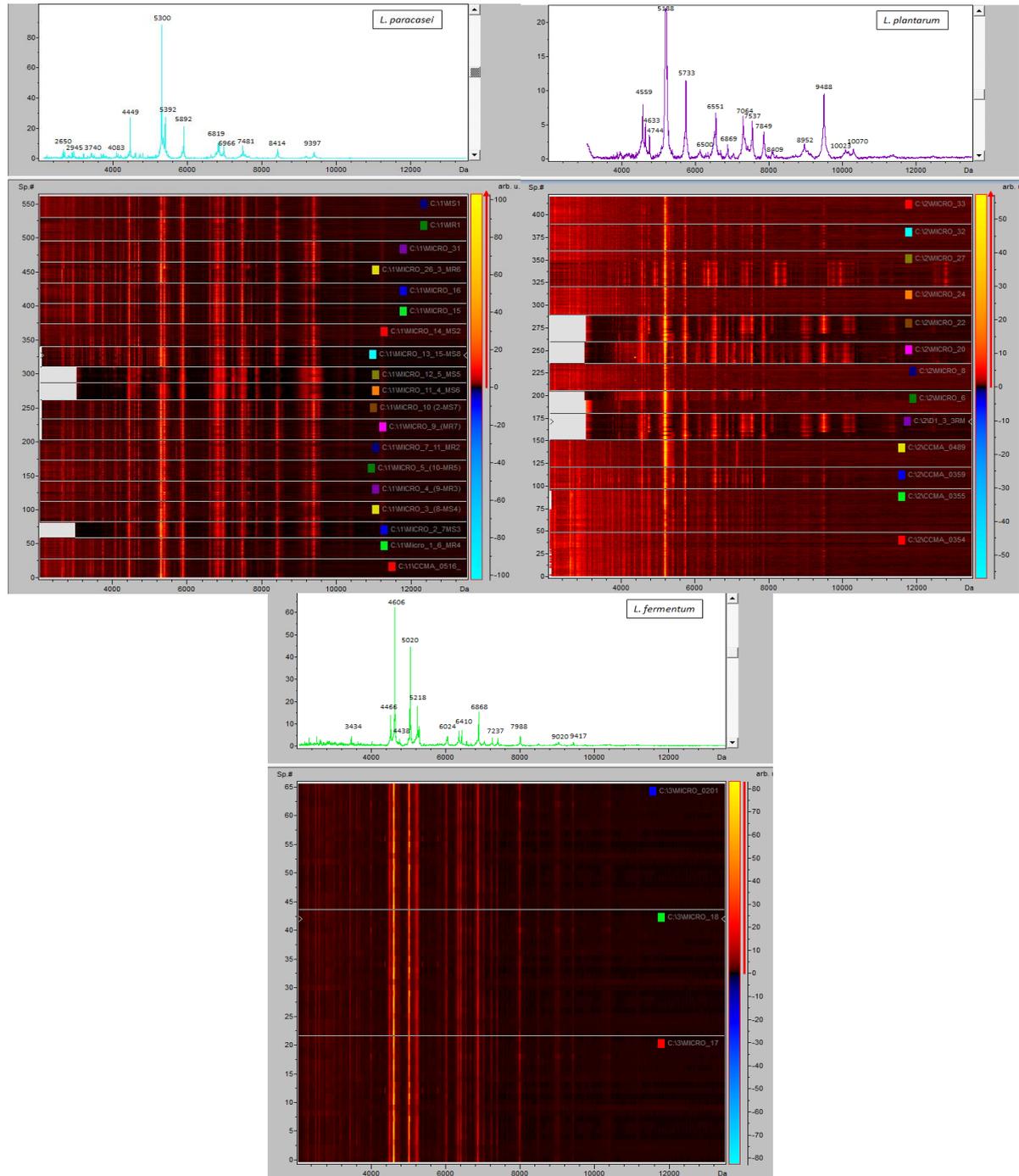
Analisando os espectros de massa classificados como *L. plantarum* (espectros numerados 19 a 25, canto superior esquerdo da Figura 16) notamos que uma característica comum a todos os isolados foi a presença de um pico  $m/z$  5188 mais intenso e característico da espécie. Ao contrário dos isolados classificados como *L. paracasei*, não houve ambiguidade na classificação dos isolados *L. plantarum* em relação à classificação baseada nos dois genes

*housekeeping*, tornando-os mais fáceis de serem identificados, apesar de, notavelmente, possuírem menor quantidade de proteínas passíveis de ionização, resultando em uma menor quantidade de picos no espectro. O mesmo ocorreu com os dois isolados classificados como *L. fermentum* (espectro numerado 26 e 27), que foi classificado como pertencente a mesma espécie pelos três genes e pelo banco de dados Biotyper, tendo dois picos mais intensos característicos com  $m/z$  4607 e 5020, respectivamente. Novamente, DUSKOVÁ et al. (2012) obteve resultados semelhantes em relação a classificação dos isolados pertencentes às espécies *L. plantarum* e *L. fermentum*, as quais tiveram a mesma classificação tanto pelo gene 16S rDNA quanto pelo Biotyper em 86% e 75% das vezes. Desta forma, analisando os resultados de maneira geral, 97% dos *Lactobacillus* tiveram a mesma classificação que o gene *pheS* e 81% que o gene *groEL*, diferenciando apenas em relação aos isolados classificados como *L. casei* pelos genes *housekeeping*.

Para visualizar mais detalhadamente quais regiões espectrais eram constantes e quais eram variáveis e, se essa característica era encontrada em todos os indivíduos da mesma espécie e, portanto, consideradas regiões intra-específicas, os isolados foram primeiramente separados em grupos de acordo com sua classificação molecular. Desta forma, 19 isolados formaram o grupo dos *L. paracasei*; 13 isolados do grupo *L. plantarum*; 3 do *L. fermentum*; e 1 do *L. brevis*. Os arquivos provindos do MALDI-TOF MS, das 30 repetições de cada isolado de cada grupo foram visualizados no software ClinProTools (Figura 17).

Na Figura 17, são mostradas duas imagens que são geradas concomitantemente no software: 1) um perfil espectral relacionando a intensidade dos picos em unidades arbitrárias (eixo das ordenadas) e o valor de  $m/z$  (eixo das abscisas); 2) uma imagem que exhibe todos os espectros carregados de todas as classes organizadas em uma aparência de pseudo-gel e cada classe é separada da próxima linha carregada por uma linha horizontal. O eixo das abscissas registra o valor  $m/z$  e o eixo das ordenadas do lado esquerdo exhibe o número do espectro em execução originado do carregamento de espectros subsequente. A intensidade do pico é expressa por um código de cores. A barra de cores e o eixo das ordenadas do lado direito indicam a relação entre a cor com a qual um pico é exibido e a intensidade do pico em unidades arbitrárias. Desta forma, analisando os espectros com suas respectivas visualizações em forma de pseudo-gel, nota-se que houve picos que permaneceram constantes em todas as repetições de todos os isolados pertencentes a mesma espécie, quando foram crescidos no mesmo meio seletivo, MRS com ágar. Adicionalmente, nota-se que os picos mais intensos foram também os mais característicos da espécie, havendo uma correlação positiva entre a intensidade do pico e a presença do mesmo em todos os isolados do grupo, intensificando a possibilidade desses picos serem utilizados como biomarcadores intraespecíficos. Para evidenciar quais seriam exatamente esses picos, os valores de  $m/z$  que apareceram com 100% de frequência nas repetições de todos os isolados de cada grupo estão listados na tabela 4.

A existência de valores de  $m/z$  presentes em todos os isolados de uma espécie e, ainda, com alta intensidade, proporciona a possibilidade desses picos serem usados como marcadores interespecíficos, devido a sua facilidade de ionização e constância, entretanto, para isso, teria que ser avaliado em um acervo muito maior de *Lactobacillus*. Contudo, de uma maneira geral, a identificação microbiológica por meio de perfil espectral por MALDI-TOF MS mostrou-se uma técnica de identificação consistente, reprodutível e fidedigna, com potencial para encontrar possíveis biomarcadores intraespecíficos.



**Figura 17.** Regiões espectrais comuns de todos isolados de *Lactobacillus* crescidos em meio seletivo MRS com ágar, pertencentes a três espécies distintas: *L. paracasei*, *L. plantarum* e *L. fermentum*, utilizando o software *ClinProTools*.

\*Parte superior: perfil espectral representativo da espécie relacionando a intensidade do pico (unidades arbitrárias, eixo das ordenadas) e os valores de  $m/z$  (eixo das abcissas). Parte inferior: pseudo-gel com todas as repetições carregadas de cada microrganismo que relaciona os valores de  $m/z$  (eixo das abcissas), número do espectro em execução (eixo das ordenadas esquerdo) e a intensidade do pico de acordo com a escala de cores (eixo das ordenadas direito).

**Tabela 4.** Valores de m/z presentes em 100% das repetições de todos os isolados de cada grupo de espécies.

<i>m/z</i> com 100% de frequência em cada grupo		
<i>L.paracasei</i>	<i>L.plantarum</i>	<i>L.fermentum</i>
5300*	5185*	3346
5890	5222	3441
6752	5733	4607*
6846		4718
6965		6564
7480		6972
8412		8141
9395		9057
		9280
		9432

\*o asterisco (\*) ao lado do número significa que foi o m/z com maior intensidade do grupo.

#### **6.3.4. Comparação do poder discriminatório para identificação de isolados de *Lactobacillus* ao nível de espécie entre o banco de dados do Biotyper e o banco de dados interno desenvolvido**

Para verificar se haveria um aumento substancial na percentagem dos isolados que seriam identificados ao nível de espécie com a adição dos espectros de microrganismos do processo de fermentação etanólica ao banco de dados, foram comparados os scores correspondentes de todos os microrganismos em relação ao banco de dados oficial do Biotyper e em relação ao banco de dados interno criado neste estudo. Os resultados das comparações podem ser visualizados na Tabela 5.

Nota-se que 51% do total dos microrganismos não foram identificados ao nível de espécie por nenhuma repetição técnica ou biológica quando foi utilizado apenas o banco de dados Biotyper como referência, enquanto que utilizando os mesmos microrganismos contra o banco de dados interno criado neste trabalho, 97% dos microrganismos tiveram alguma percentagem na identificação das suas replicatas técnicas e biológicas ao nível de espécie, sendo que 86% dos microrganismos tiveram o score acima de 2.3 em 70% das repetições e 22% tiveram score acima de 2.3 em 100% das repetições. Além disso, 17% dos microrganismos tiveram alguma repetição com score insuficiente para qualquer tipo de classificação, tornando a classificação inválida quando foi utilizado o banco Biotyper, enquanto que utilizando o banco de dados interno, nenhum microrganismo obteve classificação inválida.

Entretanto, 51% dos microrganismos foram classificados ao nível de gênero com confiabilidade (score entre 2 e 2.299) em mais de 60% das suas repetições quando utilizado o banco de dados Biotyper, evidenciando que as impressões digitais ao nível de gênero são mais consistentes do que ao nível de espécie quando se utiliza um banco de dados com isolados clínicos. De maneira geral, 22 dos 35 microrganismos testados tiveram um aumento de 57 a 100% das repetições identificadas ao nível de espécie quando utilizado o banco de dados interno.

**Tabela 5.** Porcentagem das repetições técnicas e biológicas dos isolados comerciais e isolados do processo de fermentação etanólica que foram classificados de acordo com o score de similaridade espectral entre o banco de dados oficial do Biotyper e o banco de dados interno desenvolvido neste estudo. A classificação é dividida em 4 categorias pelo próprio software: 1) Scores acima de 2.300 = alta probabilidade de identificação ao nível de espécie; 2) Score entre 2 e 2.299 = Identificação segura ao nível de gênero e provável de espécie; 3) Score entre 1.700 e 1.999 Identificação provável do gênero; 4) Score menor que 1.700 = identificação não válida

	BANCO DE DADOS BIOTYPER				BANCO DE DADOS INTERNO			
	IDENT. ESPÉCIE	IDENT. GENERO	PROV. IDENT GENERO	IDENT. INVÁLIDA	IDENT. ESPÉCIE	IDENT. GENERO	PROV. IDENT GENERO	IDENT. INVÁLIDA
CCMA_0354	0%	6%	79%	15%	94%	6%	0%	0%
CCMA_0355	0%	25%	75%	0%	98%	2%	0%	0%
CCMA_0359	45,8%	52%	4%	0%	79%	20%	0%	0%
CCMA_0383	23%	74%	3%	0%	ausente	ausente	ausente	ausente
CCMA_0489	0%	70%	20%	14%	20%	70%	10%	0%
CCMA_0516	58%	38%	4%	0%	96%	4%	0%	0%
CCMA_0201	0%	100%	0%	0%	100%	0%	0%	0%
MICRO_1	16%	77%	6%	0%	97%	3%	0%	0%
MICRO_2	13%	79%	8%	0%	71%	25%	4%	0%
MICRO_3	50%	50%	0%	0%	97%	3%	0%	0%
MICRO_4	37%	60%	3%	0%	100%	0%	0%	0%
MICRO_5	3%	97%	0%	0%	97%	3%	0%	0%
MICRO_6	0%	0%	90%	10%	60%	40%	0%	0%
Micro_7	53%	33%	13%	0%	90%	10%	0%	0%
MICRO_8	0%	10%	54%	36%	64%	25%	11%	0%
MICRO_9	71%	26%	3%	0%	98%	2%	0%	0%
MICRO_10	96%	4%	0%	0%	100%	0%	0%	0%
MICRO_11	0%	88%	12%	0%	88%	12%	0%	0%
MICRO_12	0%	75%	25%	0%	71%	29%	0%	0%
Micro_13	97%	3%	0%	0%	100%	0%	0%	0%
Micro_14	61%	39%	0%	0%	97%	3%	0%	0%
Micro_15	24%	73%	3%	0%	97%	3%	0%	0%
Micro_16	24%	73%	3%	0%	97%	3%	0%	0%
Micro_17	0%	100%	0%	0%	100%	0%	0%	0%
Micro_18	0%	100%	0%	0%	100%	0%	0%	0%
Micro_20	0%	63%	37%	0%	0%	79%	21%	0%
Micro_22	0%	80%	20%	0%	70%	27%	3%	0%
Micro_24	0%	26%	45%	29%	71%	23%	6%	0%
Micro_26	65%	35%	0%	0%	100%	0%	0%	0%
Micro_27	0%	3%	97%	0%	53%	47%	0%	0%
Micro_31	64%	36%	0%	0%	100%	0%	0%	0%
Micro_32	0%	57%	43%	0%	79%	21%	0%	0%
Micro_33	0%	9%	59%	32%	25%	50%	25%	0%
MR1	0%	97%	3%	0%	84%	16%	0%	0%
MS1	34%	60%	6%	0%	91%	9%	0%	0%
D1_3_3RM	0%	96%	4%	0%	84%	16%	0%	0%

Ressalta-se que toda vez que era utilizado o banco de dados interno, o microrganismo a ser identificado era retirado do banco antes da identificação, desta forma não ocorria a sobreposição com os espectros do próprio microrganismo resultando num score muito alto, evitando, assim, uma identificação falso-positiva. Além disso, o isolado denominado CCMA\_0383 foi retirado para o cálculo das porcentagens, pois não havia nenhum outro isolado pertencente a espécie *L. brevis* no banco de dados interno que possibilitasse a identificação do mesmo, restringindo a identificação apenas utilizando o banco de dados oficial do Biotyper.

A partir dos dados mostrados na Tabela 5, claramente houve um aumento substancial do poder discriminatório interespecífico com a mudança do banco de dados oficial do Biotyper para um banco de dados específico de microrganismos agrários brasileiros, melhorando a confiabilidade das identificações. Nota-se que a identificação ao nível de gênero é mais robusta, sendo eficaz em grande parte dos isolados utilizando um banco de dados não-específico. Entretanto, para identificação ao nível de espécie, a discriminação espectral é mais sutil, sendo necessário o desenvolvimento de um banco interno de isolados do mesmo ambiente. Estes resultados mostram a importância de construir um banco de dados próprios da biodiversidade microbiana encontrada nas dornas de fabricação de etanol brasileiro para a identificação rápida e precisa destes microrganismos.

### 6.3.5. Análise estatística multivariada de isolados de *Lactobacillus* a partir dos dados de MALDI-TOF MS

Para verificar se as diferenças espectrais interespecíficas e/ou as semelhanças intraespecíficas encontradas nos perfis espectrais eram também suficientes para distinguir estatisticamente os isolados em diferentes grupos, e, desta forma, validar a eficiência discriminatória da técnica, foram realizadas análises estatísticas multivariadas utilizando o conjunto de ferramentas online do *Metabo-Analyst* (versão 4.0), a partir de uma matriz de dados que correlacionava os valores de  $m/z$  de cada pico identificado com sua respectiva intensidade relativa. As análises estatísticas realizadas foram: análise de componentes principais (PCA) (dados não mostrados), análise de mínimos quadrados parciais discriminantes (PLS-DA) (Figura 18-A), análise da importância da variável na projeção (*VIP score*) (Figura 18-B) e análise hierárquica de agrupamentos (HCA) (Figura 18-C). Esses métodos são apropriados para analisar conjuntos de dados com alta dimensão, com o objetivo de obter uma visão generalizada dos dados (PCA) e uma modelagem classificatória dos dados (PLS-DA).

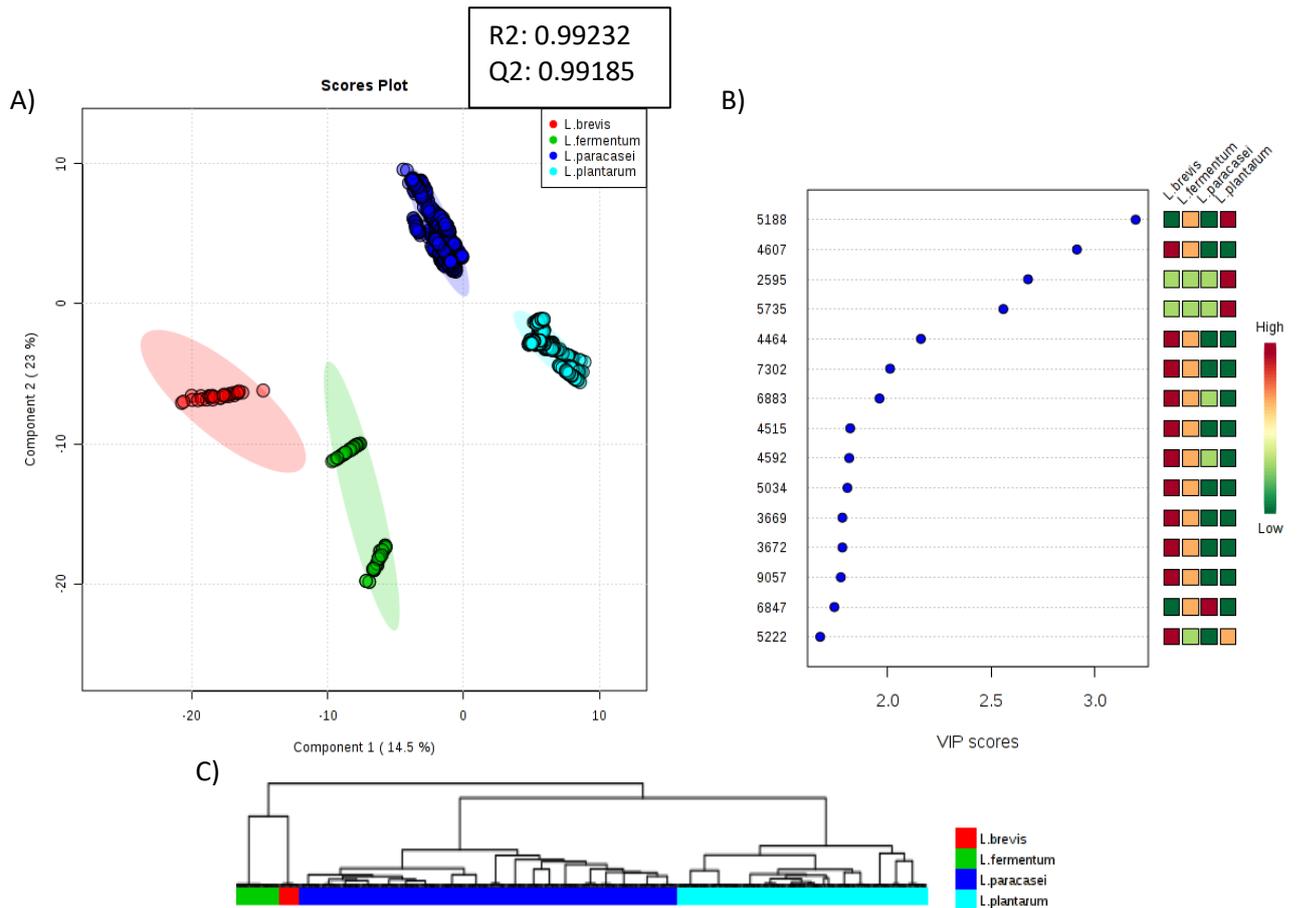
A análise de componentes principais (PCA) é um método utilizado principalmente para compressão dos dados, ou seja, reduz a alta dimensionalidade de dados multivariados em algumas dimensões capazes de capturar as principais fontes de variação dos dados, onde as variáveis originais são combinadas de maneira linear a partir de componentes principais e, conseqüentemente, dando uma visão preliminar da amostra, (HILARIO et al., 2004; HILARIO e KALOUSIS, 2008). Já a análise parcial de mínimos quadrados discriminantes (PLS-DA) é uma extensão do PCA, que desenvolve um modelo que objetiva a classificação das amostras. Essa classificação é realizada por meio de um modelo que leva em consideração a relação entre duas matrizes, uma contendo a intensidade relativa de cada pico do espectro com os valores de  $m/z$  e outra contendo a filiação de classes para cada amostra. O algoritmo decompõe as duas matrizes utilizando um número determinado de variáveis latentes, de forma que se obtenha a máxima separação entre as classes (MORS et al., 2000; SCHIOZER et al., 2012). Além disso, o PLS-DA também realiza a validação cruzada, que é a retirada de um subconjunto de amostras do conjunto de calibração e construção do modelo utilizando apenas as amostras restantes e, subsequente aplicação do modelo utilizando as amostras retidas

(NAES et al., 2002). Outro parâmetro calculado importante é o coeficiente de determinação ( $R^2$ ) do modelo proposto, ou seja, é a medida que define quão ajustado o modelo está em relação a amostra, representando a qualidade do modelo. Os valores de  $R^2$  são entre 0 e 1, sendo a qualidade do modelo mais alta quando o valor fica mais próximo de 1. Mais um parâmetro de validação da discriminação do modelo é o  $Q^2$ , o qual mostra a potência de predição do modelo.

A modelagem PLS-DA têm sido utilizada em várias áreas da proteômica (LEE et al., 2003; NORDEN et al., 2005; LILEY e CUPREE, 2006; PIERCE et al., 2006; FENG et al., 2011). Estudos proteômicos realizados por Lee e colaboradores (2003) utilizaram a modelagem PLS-DA em dados de espectrometria de massas com 60.000 variáveis  $m/z$  e reduziram o conjunto de dados para 545 variáveis  $m/z$ , identificando biomarcadores que potencialmente contribuíram para a discriminação de amostras saudáveis e cancerígenas.

Desta forma, analisando a PLS-DA realizado com os dados deste trabalho (Figura 18-A) nota-se que os dois primeiros componentes principais explicaram mais de 37.5% da separação dos grupos, sendo necessário 5 componentes para explicar 68% da separação, permitindo a separação dos isolados em 4 grupos distintos e não-sobrepostos, identificados por 4 cores diferentes: *L. paracasei* (azul escuro), sendo o grupo com maior quantidade de isolados; *L. plantarum* (azul claro), sendo o segundo com maior número de isolados; *L. fermentum* (verde); e por último o *L. brevis* (vermelho), com apenas 1 isolado e 30 repetições. A mesma classificação de cores e número de isolados foi utilizada para construir o dendograma a partir da análise hierárquica de agrupamentos (Figura 18- C). Notavelmente, os isolados de *Lactobacillus* também foram subdivididos em 4 grupos distintos, não havendo sobreposição sobre os mesmos por nenhuma repetição. Contudo, houve uma divisão prévia em 2 clusters principais, mostrando que os isolados de *L. fermentum* e *L. brevis* eram espectralmente mais semelhantes entre si do que os isolados de *L. paracasei* e *L. plantarum*, os quais ficaram juntos em outro cluster separado. Pelo dendograma, não podemos inferir uma proximidade evolutiva entre os isolados que ficaram no mesmo cluster, apenas podemos concluir que possuíam picos com  $m/z$  e intensidades relativas mais similares, provavelmente por compartilharem alguns picos, entretanto, a dissimilaridade entre os quatros grupos foi o suficiente para distingui-los em grupos distintos.

Já a pontuação VIP é uma medida que resume a contribuição de uma variável para a construção do modelo PLS-DA. A figura 18-B mostra os 15 valores de  $m/z$  (variáveis) que mais contribuíram para a discriminação dos grupos de acordo com a pontuação VIP (eixo das ordenadas). Nota-se que mais da metade dos  $m/z$  que mais contribuíram para o modelo são picos compartilhados apenas entre *L. brevis* (alta intensidade) e *L. fermentum* (baixa intensidade) ( $m/z$  4607, 4464, 7302, 6883, 4515, 4592, 5034, 3669, 3672 e 9057), o que provavelmente explica porque esses dois isolados ficaram na mesma chave do dendograma. Em contrapartida, os demais picos eram apenas intensos em *L. plantarum* ( $m/z$  2592, 5188, 5735), sendo característicos desse grupo.



**Figura 18.** A) Análise de mínimos quadrados parciais discriminantes (PLS-DA) utilizando valores de  $m/\lambda$  e intensidade relativa de todos os isolados de *Lactobacillus* separados em grupos de acordo com o componente 1 (eixo das abcissas, 14,5 %) e componente 2 (eixo das ordenadas, 23%) e os valores de  $Q^2$  e  $R^2$  do modelo proposto. B) Importância da variável na projeção (VIP score) utilizando os 15 valores de  $m/\lambda$  (eixo das ordenadas, lado esquerdo) que mais contribuíram para o modelo de separação PLS-DA com suas respectivas intensidades de acordo com a escala de cor, sendo vermelho referente aos picos mais intensos ou específicos do microrganismo e verde aos menos intensos ou ausentes naquele microrganismo (eixo das ordenadas, lado direito). C) Dendrograma dos isolados de *Lactobacillus* de acordo com a sua similaridade espectral.

## 6.4. Aquisição dos espectros de *Saccharomyces* e análise dos dados obtidos por MALDI-TOF MS

### 6.4.1. Perfil proteico espectral dos isolados de *Saccharomyces*

O perfil espectral obtido por MALDI-TOF MS das três leveduras *S. cerevisiae* selecionadas especificamente para o processo de fermentação e produção de etanol de usinas brasileiras e da levedura, também *S. cerevisiae*, pertencente ao banco de dados da Universidade Federal de Lavras, isolada a partir da fermentação de cacau, assim como suas regiões espectrais comuns, podem ser visualizados na Figura 19. Ressalta-se que todos os isolados foram crescidos no meio seletivo YEPD, sob as mesmas condições de cultivo, preparo e aplicação de amostra na placa MALDI.

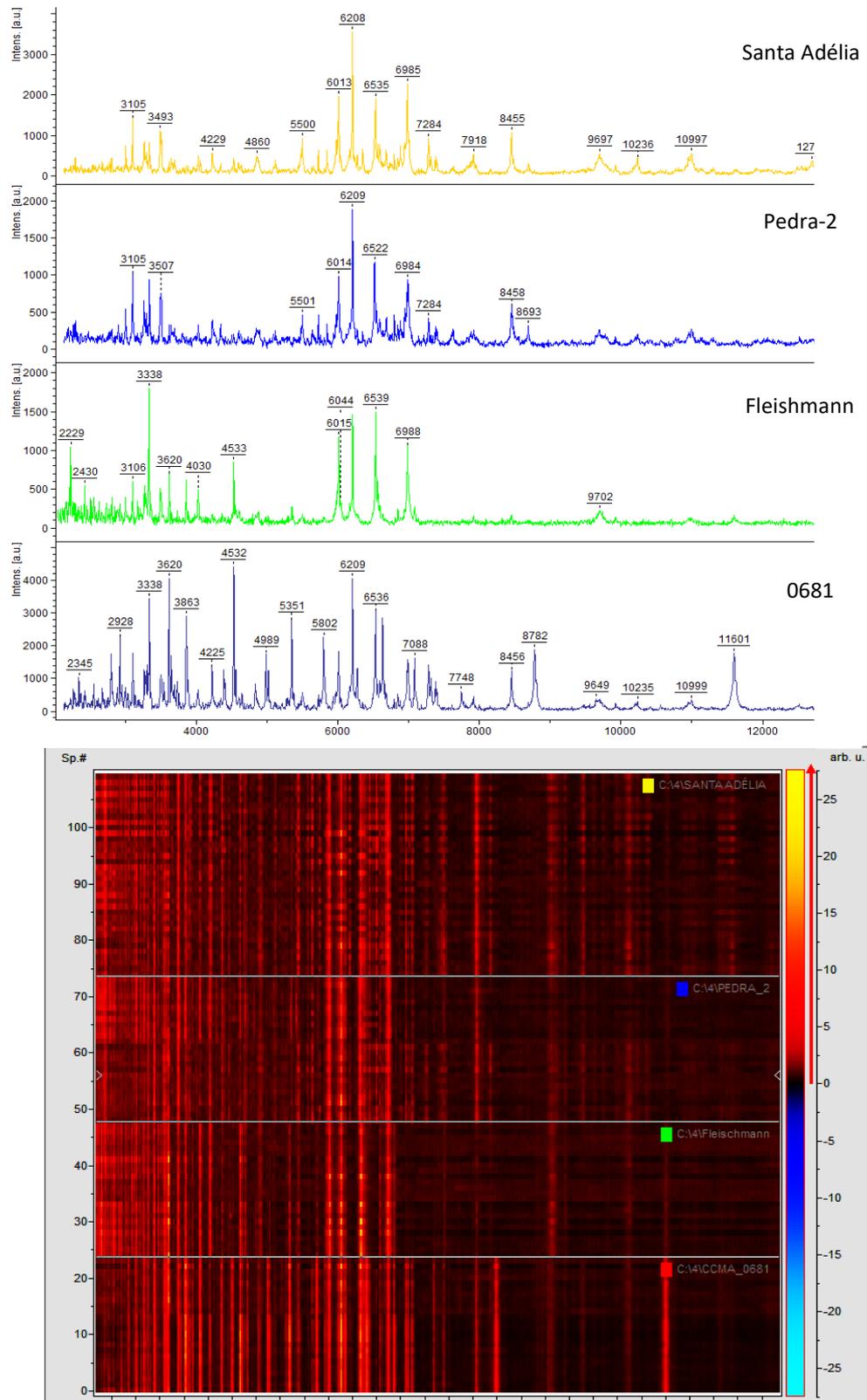
Correlacionando o perfil espectral obtido pelo software Flex Analysis (3.0, Bruker) com a figura de pseudogel que relaciona a intensidade dos picos a intensidade de cor adquirido pelo software ClinProTools (3.0,

Bruker) apresentadas juntas na Figura 19, nota-se que os espectros de massa das leveduras forneceram impressões espectrais semelhantes, com picos compartilhados entre os quatro isolados, tais como  $m/z$  6015, 6210, 6537 e 6987, provavelmente característicos da espécie. Entretanto, o isolado 0681 possuiu uma maior diversidade e quantidade de picos, enquanto os isolados Santa-Adélia, Fleishmann e Pedra-2 possuíram perfis mais similares. Para verificar se as diferenças nas assinaturas espectrais encontradas eram suficientes para discriminar os isolados em diferentes grupos, foram feitas análises estatísticas multivariadas no Metabolytic a partir de uma matriz de dados com todas as repetições, que correlacionava os valores de  $m/z$  dos picos (com uma tolerância de 10 daltons) com suas respectivas intensidades relativas. As análises realizadas foram: análise de componentes principais (PCA) (dados não mostrados), análise de mínimos quadrados parciais discriminantes (PLS-DA) (Figura 20-A), análise da importância da variável na projeção (VIP score) (Figura 20-B) e análise hierárquica de agrupamentos (HCA) acoplada à análise de expressão diferencial de acordo com a intensidade dos íons (Heat Map) (Figura 20-C).

Analisando a figura 20 como um todo, temos que a análise estatística PLS-DA com apenas dois componentes (Componente 1 no eixo das abcissas e Componente 2 no eixo das ordenadas) explicou 74% da separação dos grupos, sendo necessários 5 componentes para explicar 100% da separação dos 4 isolados de *S.cerevisiae* em diferentes grupos não-sobrepostos (ilustrados de diferentes cores na legenda da Figura 20-A). Na figura 20-B, temos os 15 valores de  $m/z$  que mais contribuíram para a construção do modelo PLS-DA de acordo com a pontuação VIP (eixo das ordenadas) e suas intensidades relativas presentes em cada isolado. Nota-se que mais de 50% dos  $m/z$  selecionados ( $m/z$  3619, 6634, 4989, 3647, 5351, 11.599, 5801 e 4224) pela análise de VIP score, foram mais intensos no isolado 0681, evidenciando ser o isolado mais discrepante dos quatro analisados. Em relação aos mesmos picos, houve uma padronização em relação aos outros isolados, sendo sempre mais intenso no isolado PEDRA-2, menos intenso em SANTA-ADÉLIA e praticamente inexistente em FLEISHMANN. Outra característica notável foi a existência de 4 picos apenas nos isolados da fermentação etanólica,  $m/z$  6986, 3494, 6948 e 6800, sendo sempre mais intenso no isolado SANTA-ADÉLIA.

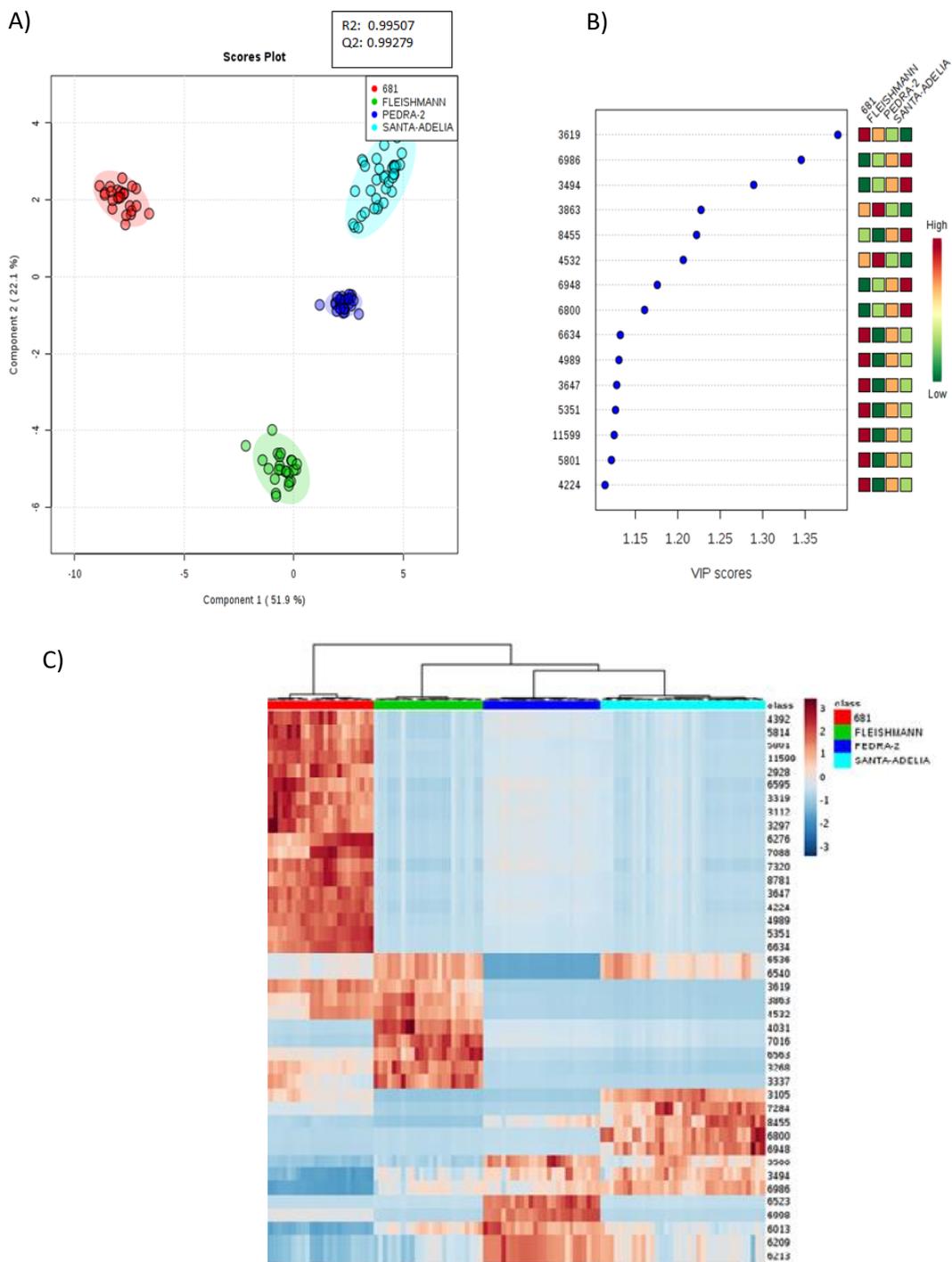
Já na figura 20-C, podemos observar as diferentes intensidades expressas de todos os valores de  $m/z$  individualmente, comparando os 4 isolados e suas distribuições no dendograma. Nitidamente percebemos que o dendograma separou os isolados em 4 grupos distintos, não sobrepostos e que o isolado 0681 ficou em uma chave separada dos demais isolados da fermentação etanólica, e que os isolados SANTA-ADÉLIA e PEDRA-2 ficaram mais próximos. Essa separação é mais nítida visualizando o perfil de intensidade iônica no Heat Map, em que claramente percebe-se que cada isolado possui um perfil próprio, sendo os perfis dos isolados 0681 e FLEISHMANN mais discrepantes e singulares, e os perfis de SANTA-ADELIA e PEDRA-2 mais próximos.

Estes resultados preliminares mostram que a técnica de identificação por MALDI-TOF para leveduras pode ser discriminatória mesmo com indivíduos pertencentes a mesma espécie, evidenciando uma potencial ferramenta para discriminação de isolados de *Saccharomyces* selecionados para o processo de produção de etanol de isolados de *Saccharomyces* que não foram selecionados para o processo e, além disso, uma técnica promissora para a distinção de cepas não-*Saccharomyces*, contaminantes do processo.



**Figura 19.** Perfil espectral dos isolados de *Saccharomyces cerevisiae*, crescidos em meio seletivo YEPD, obtido por MALDI-TOF e suas respectivas regiões espectrais compartilhadas.

\*Parte superior: perfil espectral representativo da espécie relacionando a intensidade do pico (unidades arbitrárias, eixo das ordenadas) e os valores de  $m/z$  (eixo das abscissas) pelo software FlexAnalysis (3.0, Bruker). Parte inferior: pseudo-gel com todas as repetições carregadas de cada microrganismo que relaciona os valores de  $m/z$  (eixo das abscissas), número do espectro em execução (eixo das ordenadas lado esquerdo) e a intensidade do pico de acordo com a escala de cores (eixo das ordenadas lado direito) pelo software ClinProTools (3.0, Bruker).

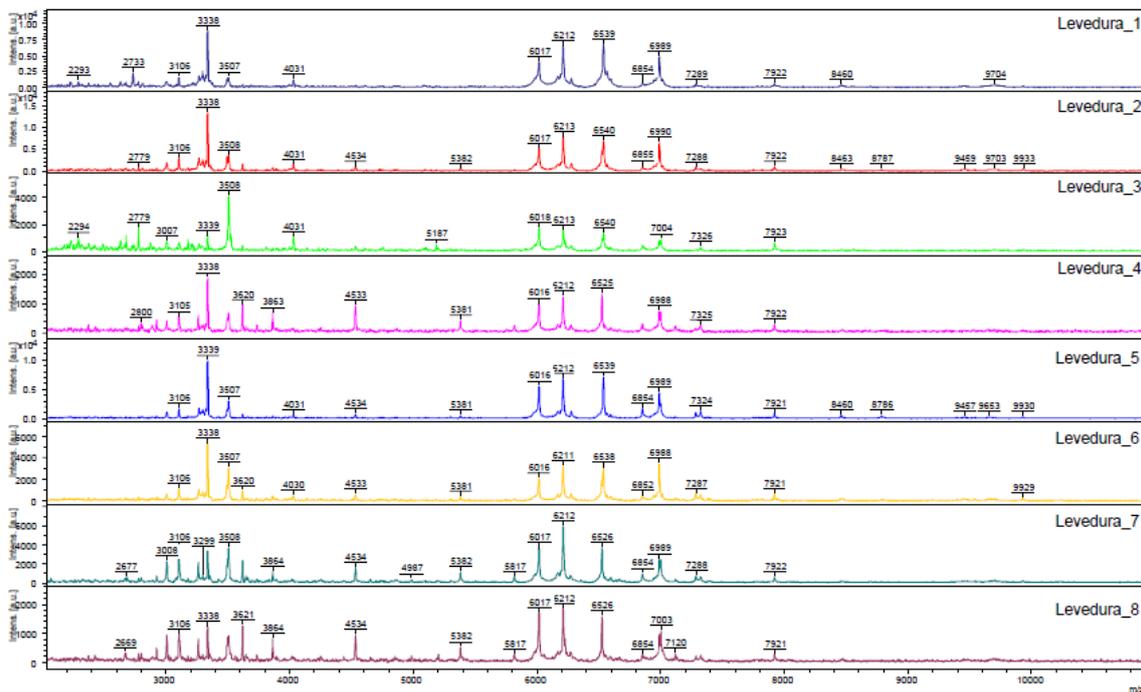


**Figura 20.** A) Análise de mínimos quadrados parciais discriminantes (PLS-DA) utilizando valores de  $m/z$  e intensidade relativa de todos os isolados de *Lactobacillus* separados em grupos de acordo com o componente 1 (eixo das abcissas, 51.9 %) e componente 2 (eixo das ordenadas, 22.1%) e os valores de  $Q^2$  e  $R^2$  do modelo proposto. B) Importância da variável na projeção (VIP score) utilizando os 15 valores de  $m/z$  (eixo das ordenadas, lado esquerdo) que mais contribuíram para o modelo de separação PLS-DA com suas respectivas intensidades de acordo com a escala de cor, sendo vermelho referente aos picos mais intensos ou específicos do microrganismo e verde aos menos intensos ou ausentes naquele microrganismo (eixo das ordenadas, lado direito). C) Dendrograma dos isolados de *Lactobacillus* de acordo com a sua similaridade espectral e o Heat Map que correlaciona os valores de  $m/z$  com uma escala de cor de acordo com a intensidade do pico relatada em cada isolado.

Investigando a capacidade da técnica de MALDI-TOF MS para biotipagem de cepas de leveduras, selecionadas e contaminantes, na indústria, foi realizado um teste preliminar a partir de amostra coletada diretamente da dorna em escala industrial e em condições ótimas de fermentação. Após serem feitos diluições

seriadas em meio de cultura para obtenção de colônias individuais com antibiótico, 8 colônias foram submetidas ao preparo e aplicação de amostra para obtenção do perfil espectral proteico por MALDI-TOF MS. Posteriormente, os dados espectrais brutos foram comparados com todo acervo do banco de dados Biotyper com 7.311 perfis espectrais proteicos depositados, entre eles, 13 espectros de *Saccharomyces cerevisiae*, 2 espectros de *Dekkera* e 77 espectros de *Pichia*, e espectros de outros microrganismos considerados leveduras contaminantes do processo de produção de etanol, juntamente com os espectros dos quatro isolados deste estudo, pertencente ao banco interno complementar. Os perfis espectrais e suas respectivas classificações e notas obtidas pelo MALDI-Biotyper das oito leveduras podem ser visualizadas na Figura 21 A-B.

A)



B)

	Log (Score) Biotyper	Classificação_Biotyper
Levedura_1	2.457	<i>S. cerevisiae</i> _Fleishmann
Levedura_2	2.436	<i>S.cerevisiae</i> _Pedra-2
Levedura_3	2.336	<i>S.cerevisiae</i> _Fleishmann
Levedura_4	2.38	<i>S.cerevisiae</i> _Pedra-2
Levedura_5	2.405	<i>S.cerevisiae</i> _Pedra-2
Levedura_6	2.454	<i>S.cerevisiae</i> _Pedra-2
Levedura_7	2.427	<i>S.cerevisiae</i> _Pedra-2
Levedura_8	2.41	<i>S.cerevisiae</i> _Pedra-2

**Figura 21.** A) Perfil espectral proteico das oito leveduras isoladas diretamente do processo de produção de etanol por diluição seriada utilizando MALDI-TOF MS; B) e suas respectivas notas e classificações utilizando o banco de dados oficial Biotyper, juntamente com o banco de dados complementar desenvolvido neste estudo.

\*Ordenada: intensidade relativa; abscissa: massa / carga ( $m/z$ ); Levedura\_n: identificação das colônias de leveduras isoladas; Log (score) Biotyper: nota da simetria espectral dos microrganismos desconhecidos em relação aos espectros do banco de dados.

Os resultados de log(score) do Biotyper mostraram que as 8 colônias tiveram a maior simetria espectral com os espectros de leveduras do banco de dados interno e puderam ser diferenciadas em duas diferentes cepas,

PEDRA-2 e FLEISHMANN, com um score acima de 2.3 (identificação confiável ao nível de espécie). Desta forma, os resultados revelaram que apesar de ter um acervo de *S.cerevisiae* maior no banco de dados oficial do Biotyper (13 isolados) em relação ao banco de dados complementar (apenas 4 isolados), o log score foi maior quando comparado com os isolados de *S.cerevisiae* do próprio processo. Estudos realizados por Moothoo-Padayachie e colaboradores (2013) para biotipagem de cepas comerciais de *S.cerevisiae* por MALDI-TOF MS, também testaram o poder de biotipagem da técnica para a indústria de produção de vinhos, cervejas e etanol. Para isso, primeiramente, 20 cepas providas do Conselho de Pesquisa Agrícola (*Agricultural Research Council – ARC*), da África do Sul, tiveram seus códigos de identificação apagados e os seus perfis espectrais foram contrastados com os espectros de 44 isolados de *Saccharomyces cerevisiae* pertencentes ao banco de dados complementar desenvolvido pelo estudo dentro do Biotyper. Os 20 microrganismos foram classificados corretamente ao nível de gênero e espécie e 18 foram classificados corretamente ao nível de estirpe, portanto, MALDI-TOF MS foi capaz de identificar isolados de *S. cerevisiae* ao nível de gênero e espécie com 100% de precisão e ao nível de estirpe com 90% de precisão. Posteriormente, em outro teste, 7 isolados de leveduras foram obtidos de uma planta produtora de etanol industrial local e os perfis espectrais obtidos foram também contrastados com o banco interno no Biotyper. Destas 7 cepas, apenas 4 foram identificadas corretamente, correspondendo às culturas de creme levedura puras. Já as 3 cepas incorretamente biotipadas, foram recuperadas de três tanques de fermentação problemáticos (com baixo rendimento de álcool), não sendo culturas puras e, ficando agrupadas separadamente na análise estatística PCA. O estudo sugere que o baixo rendimento de etanol pode ter sido causado pela contaminação da cultura pura e que a técnica MALDI-TOF MS mostrou ter um grande potencial como ferramenta de identificação rápida de linhagens industriais de *S. cerevisiae* e em breve será aplicada para garantir a qualidade das culturas de leveduras e detectar contaminações que poderiam levar a grandes perdas em larga escala.

## 7. CONSIDERAÇÕES FINAIS

Este estudo permitiu, primeiramente, estabelecer duas metodologias de extração proteica, uma específica para bactérias ácido lácticas e outra para leveduras. Essas abordagens levaram em consideração as características físico-químicas dos microrganismos estudados. Os diferentes métodos de aplicação e aquisição dos espectros desses dois microrganismos também foram avaliados, sempre visando a rápida identificação por MALDI-TOF MS. Em outro pilar do projeto, parâmetros críticos da técnica, como concentração celular e crescimento microbiano em diferentes meios de cultura foram ponderados, verificando que mesmo havendo algumas variações no perfil espectral de acordo com o meio de cultivo, essas diferenças não foram suficientes para uma classificação errônea pelo banco de dados MALDI-Biotyper.

Assim como toda técnica baseada em espectrometria de massas, a biotipagem por MALDI-TOF MS possui certas limitações. Dentre as principais estão: fenômenos de extinção de íons, em que o sinal baixo de algumas proteínas é extinto por íons altamente ionizáveis, suprimindo o sinal, principalmente pela intensidade ser relativa (WILKINS E LAY, 2006), podendo resultar na diminuição da sensibilidade quando estão sendo analisadas diferentes estirpes, altamente relacionadas (MOOTHOO-PADAYACHIE et al., 2012); diferentes habilidades das proteínas em formar íons, modificando a intensidade do sinal; e, principalmente, ser dependente de um banco de dados.

Nosso estudo mostrou que 63% dos microrganismos testados, pertencentes ao gênero *Lactobacillus*, tiveram um aumento da capacidade de identificação ao nível de espécie de 57% a 100%, quando foi utilizado um banco de dados próprio, com isolados do próprio processo, desta forma, ficou evidente a necessidade de construir um banco de dados cada vez mais abrangente com a biodiversidade das dornas. Entretanto, esse aumento da performance só foi possível devido a prévia caracterização molecular dos isolados utilizando técnica multilocus. Para *Lactobacillus*, a região 16SrDNA sozinha não foi discriminatória o suficiente, entretanto a utilização de genes *housekeeping* mostrou-se eficiente para a identificação interespecífica dos isolados, só não conseguindo diferenciar consistentemente alguns isolados pertencentes ao grupo *L.casei*, havendo ambiguidade na classificação quando utilizava o gene *groEL*. Já em relação as leveduras, as regiões ITS e 28S nr-LSU remeteram a mesma espécie, provendo uma maior confiabilidade classificatória. Comparando as classificações moleculares e espectrométricas de bactérias e leveduras, tivemos que a análise de *fingerprinting* por MALDI-TOF MS utilizando o banco Biotyper teve desempenho equivalente a análise MLST, mas possuindo duas vantagens principais: ser muito mais rápida, não precisando de extensas metodologias de extração, sequenciamento e montagem de seqüências e, ser mais barata.

Adicionalmente, este estudo permitiu a eficiente discriminação por análises estatísticas dos isolados de *Lactobacillus* em diferentes espécies e enumerou os valores de  $m/z$  presentes em 100% das repetições de cada grupo, como sendo prováveis picos biomarcadores intraespecíficos. Para leveduras, a análise de *fingerprinting* teve um poder discriminatório ainda maior, diferenciando cepas da mesma espécie e, mostrando que as leveduras do processo fermentativo possuíam um perfil espectral mais semelhante do que a cepa de outro ambiente, podendo ser facilmente distinguida.

Desta forma, foi demonstrado que a técnica MALDI-TOF MS possui um enorme potencial para ser utilizada como ferramenta de biotipagem de isolados de *Lactobacillus* e cepas de *S.cerevisiae*, sendo necessário a construção de um banco de dados que abranja o máximo da biodiversidade microbiológica das usinas do Brasil e, consequentemente melhore a capacidade discriminatória da técnica. Futuros estudos inter-laboratoriais podem levar ao desenvolvimento desse banco e consequentemente, a identificação rápida e precisa de cepas selecionadas de

*S.cerevisiae*, distinção de cepas prejudiciais ao processo ditas não- Saccharomyces e identificação de microrganismos bacterianos contaminantes.

## REFERÊNCIAS

- ABREU-CAVALHEIRO A, MONTEIRO G. Solving ethanol production problems with genetically modified yeast strains. **Braz J Microbiol.**, v.4, n.3, p.665–71, 2014.
- ALATOOM, A. A., CUNNINGHAM, S. A., IHDE, S. M., MANDREKAR, J., PATEL, R. Comparison of Direct Colony Method versus Extraction Method for Identification of Gram-Positive Cocci by Use of Bruker Biotyper Matrix-Assisted Laser Desorption Ionization–Time of Flight Mass Spectrometry. **Journal of Clinical Microbiology**, p. 2868–2873, 2011.
- ALLEN, J., DAVEY, H. M., BROADHURST, D., HEALD, J. K., ROWLAND, J. J., OLIVER, S. G., KELL, D. B. High-throughput classification of yeast mutants for functional genomics using metabolic footprinting. London: **Nature biotechnology**. 21: 692-696, 2003.
- ALLI, I., FAIRBAIRN, R., BAKER, B. E. The effects of ammonia on the fermentation of chopped sugarcane. Amsterda: **Animal Feed Science and Technology**. 9: 291-299 p. 1983.
- AMORIM, H. V., LEÃO, R. M. A experiência do Pro-álcool. Fermentação alcoólica: Ciência e Tecnologia, p. 190-191, **Fermentec**, Piracicaba, 2005.
- AMORIM, H. V., LOPES, M. L., DE CASTRO OLIVEIRA, J. V., BUCKERIDGE, M. S., GOLDMAN, G. H. Scientific challenges of bioethanol production in Brazil. **Appl Microbiol Biotechnol.** ;91:1267–75, 2011.
- AMORIM, H. V. DE. Infecção na fermentação: Como evitá-la. **Fermentec**, Piracicaba , 2014.
- ANDRIETTA, M. G. S., OLIVEIRA, A. J., STUPIELLO, J. P. Determinação de M.I.C (concentração inibitória mínima) para cinco antimicrobianos sobre bactérias g+ isoladas da indústria brasileira de fermentação alcoólica. **STAB. Açúcar, Álcool e Subprodutos**, Piracicaba, v.13, n.6, 42-43 p. 1995.
- ANDRIETTA, S. R., ANDRIETTA, M. G. S., BICUDO, M. P. Comparação do rendimento fermentativo utilizando diferentes metodologias de cálculo para avaliação do desempenho de um processo industrial. **STAB**, v.4, p. 41-49, 2012.
- ARAÚJO, W.L., LIMA, A.O.S., AZEVEDO, J.L., MARCON, J., KUKLINSKYSOBRAL, J., LACAVAL, P.T. Manual: isolamento de microrganismos endofíticos. **Piracicaba: CALQ**, p.86, 2002.
- ARBEFEVILLE, S., HARRIS, A., FERRIERI, P. Comparison of sequencing the D2 region of the large subunit ribosomal RNA gene (MicroSEQ®) versus the internal transcribed spacer (ITS) regions using two public databases for identification of common and uncommon clinically relevant fungal species. **J. Microbiol. Methods**, v.140, p. 40-46, 2017.
- ARNOLD, R., KARTY, J., ELLINGTON, A., REILLY, J. Monitoring the growth of a bacteria culture by MALDI-MS of whole cells. **Anal. Chem.**, v.71, p.1990-1996, 1999.
- ASSIS, D. M., JULIANO, L., JULIANO, M. A. A espectrometria de massas aplicada na classificação e identificação de microrganismos. **Revista da Universidade Vale do Rio Verde, Três Corações**, v. 9, n. 2, p. 344-355, ago./dez, 2011.
- BARREIRO, J. R. *Identificação direta de microrganismos causadores de mastite por espectrometria de massas*. 2016. 88f. Tese de Doutorado. Universidade de São Paulo, Pirassununga, 2015.
- BASÍLIO, A. C. M., ARAÚJO, P. R. L., MORAIS, J. O. F., SILVA FILHO, E. A., MORAIS, M. A. J. R., SIMÕES, D. A. Detection and identification of wild yeast contaminants of the industrial fuel ethanol fermentation process. **Curr Microbiol**, v. 56, p.322-326, 2008.

- BASSO, L.C., Amorim, H.V., Oliveira, A.J., Lopes, M.L. Yeast selection for fuel ethanol production in Brazil. **FEMS Yeast Res**, 8:1155-1163, 2008.
- BASSO, T. O., GOMES, F. S., LOPES, M. L., AMORIM, H. V., EGGLESTON, G., BASSO, L. C. Homo- and Heterofermentative lactobacilli differently affect sugarcane-based fuel ethanol fermentation. **Antonie van Leeuwenhoek**. V. 105, p.169-177, 2013.
- BEALES, N. Adaptation of microorganisms to cold temperatures, weak acid preservatives, low pH, and osmotic stress: A review. **Comprehensive Reviews in Food Science and Food Safety**, v.3, n.1, p.1- 20, 2004.
- BEGEROW, D., NILSSON, H., UNTERSEHER, M., MAIER, W. Current state and perspectives of fungal DNA barcoding and rapid identification procedures. **Appl Microbiol Biotechnol.**,v.87, p.99–108, 2010.
- BENAGLI, C., DEMARTA, A., CAMINADA, A. P., ZIEGLER, D., PETRINI, O., TONOLLA, M. A rapid MALDI-TOF MS identification database at genospecies level for clinical and environmental *Aeromonas strains*. **Plos One**. 7, (10), 2012.
- BERNARDO, K., PAKULAT, N., MACHT, M., KRUT, O., SEIFERT, H., FLEER, S., HÜNGER, F., KRÖNKE, M. Identification and discrimination of *Staphylococcus aureus* strains using matrix-assisted laser desorption/ionization-time of flight mass spectrometry. **Proteomics**, v.2, p.747-753, 2002.
- BISCHOFF, K. M., LIU, S., LEATHERS, T., WORTHINGTON, R. E., RICH, J. O. Modeling Bacterial Contamination of Fuel Ethanol Fermentation, **New York: Biotechnology and Bioengineering**. 103: 117-122, 2008.
- BLAIOTTA, G., FUSCO, V., ERCOLINI, D., APONTE, M., PEPE, O., VILLANI, F. *Lactobacillus* strain diversity based on partial hsp60 gene sequences and design of PCR- restriction fragment length polymorphism assays for species identification and differentiation. **Appl. Environ. Microbiol. Washington**, v.74, p.208-215, 2008.
- BONATELLI, M. L., QUECINE, M. C., SÁ E SILVA, M., LABATE, C. A. Characterization of contaminant bacterial communities in sugarcane first-generation industrial ethanol production. **FEMS Microbiology Letters**, Volume 364, Issue 17, 15 September, 2017
- de BRUYNE, K., SCHILLINGER, U., CAROLINE, L., BOEHRINGER, B., CLEENWERCK, I., VANCANNEYT, M., DE VUYST, L., FRANZ, C. M. A. P., VANDAMME, P. *Leuconostoc bolzapfelii* sp. nov., isolated from Ethiopian coffee fermentation and assessment of sequence analysis of housekeeping genes for delineation of *Leuconostoc* species. **Int J Syst Evol Microbiol** 57, 2952–2959, 2007.
- de BRUYNE, K., CAMU, N., LEFEBVRE, K., DE VUYST, L. & VANDAMME, P. *Weissella ghanensis* sp. nov., isolated from a Ghanaian cocoa fermentation. **Int J Syst Evol Microbiol** 58, 2721–2725, 2008<sup>a</sup>.
- de BRUYNE, K., FRANZ, C. M. A. P., VANCANNEYT, M., SCHILLINGER, U., MOZZI, F., DE VALDEZ, G. F., DE VUYST, L. & VANDAMME, P. *Pediococcus argentiniensis* sp. nov. from Argentinean fermented wheat flour and identification of *Pediococcus* species by *pheS*, *rpoA* and *atpA* sequence analysis. **Int J Syst Evol Microbiol** 58, 2909–2916, 2008b.
- de BRUYNE, K., CAMU, N., DE VUYST, L. & VANDAMME, P. *Weissella fabaria* sp. nov., from a Ghanaian cocoa fermentation. **Int J Syst Evol Microbiol** 60, 1999–2005, 2010.
- CARVALHO, G.G., MONTEIRO, R.A.B. A influencia do teor de acidez e da contaminação bacteriana do mosto no rendimento fermentativo industrial para a produção de etanol. **Engineering. Faculdades Associadas de Uberaba (FAZU)**, v. 8, p. 47–54, 2011.

- CHENEBY, D., PHILIPPOT, L., HARTMANN, A., HENAULT, C., GERMON, J.C., 16S rDNA analysis for characterization of denitrifying bacteria isolated from three agricultural soils. **FEMS Microbiology Ecology**, Amsterdam, v. 34, p. 121-128, 2000.
- CHIN, P.M., INGLEDEW, W. M. Effect of lactic acid bacteria on wheat mash fermentations prepared with laboratory backset. **Enzyme Microb Technol**, v.16, p.311-317, 1994.
- CHONG, J., SOUFAN, O., LI, C., CARAUS, I., LI, S., BOURQUE, G., WISHART, D. S., XIA, J. MetaboAnalyst 4.0: towards more transparent and integrative metabolomics analysis. **Nucl. Acids Res.**, v.46, p.486-494.
- COLLINS, M.D., JONES, D., KROPPESTEDT, R.M., SCHLEIFER, K.H. Chemical studies as a guide to the classification of *Corynebacterium pyogenes* and “*Corynebacterium haemolyticum*”. **J. Gen. Microbiol.**, V.128, p.335–341, 1982.
- COMPANHIA NACIONAL DE ABASTECIMENTO –CONAB. Safra brasileira de cana-de-açúcar. 2018.
- DELLAGLIO, F., VANCANNEYT, M., ENDO, A., VANDAMME, P., FELIS, G., CASTIONI, A., FUJIMOTO, J., WATANABE, K., OKADA, S. *Lactobacillus durianis* Leisner *et al.* 2002 is a later heterotypic synonym of *Lactobacillus vaccinoferus* Kozaki and Okada 1983. **Int. J. Syst. Evol. Microbiol.** v.56, p.1721-1724, 2006.
- DELAVENNE, E., ISMAIL, R., PAWTOWSKI, A., MOUNIER, J., BARBIER, G., LE BLAY, G. Assessment of lactobacilli strains as yogurt bioprotective cultures. **Food Control**, v.30, Issue 1, p.206-213, 2013.
- EWING, B., GREEN, P. Base-calling automated sequencer traces using Phred. II. Error probabilities. *Genome Research*, v.8, p.186-194, 1998.
- FELL, J. W., BOEKHOUT, T., FONSECA, A., SCORZETTI, G., STATZELL-TALLMAN, A. Biodiversity and systematics of basidiomycetous yeasts as determined by large-subunit rDNA D1/D2 domain sequence analysis. **Int J Syst Evol Microbiol**, v.50, p.1351–1371, 2000.
- FENSELAU, C. MALDI MS and strategies for protein analysis. **Anal. Chem.** V. 69 (21), p.661A–665A, 1997.
- FENSELAU, C. DEMIREV, P. A. Characterization of intact microorganisms by MALDI mass spectrometry. **Mass Spectrom. Rev.**, 20 (4), p. 157–171, 2001.
- FERREIRA, L., SÁNCHEZ-JUANES, F., GONZÁLES-ÁVILA, M., CEMBRERO-FUCINOS, D., HERRERO-HERNANDEZ, A., GONZÁLEZ-BUITRAGO, J. M., MUNOZ-BELLIDO, J. L. Direct identification of urinary tract pathogens from urine samples by matrix-assisted laser desorption ionization- time of flight mass spectrometry. **Journal of Clinical Microbiology**, v.48, p. 2110-2115, 2010.
- FLEET, G. H. Wine yeasts for the future. **FEMS Yeast Research**. v. 8, n.7, p. 979- 995, 2008.
- FÖRSTER, J, FAMILI, I, FU, P, PALSSON, B. Ø, NIELSEN, J. Genome-Scale Reconstruction of the *Saccharomyces cerevisiae*. **Metabolic Network.**, p. 244 - 253, 2003.
- Genome-Scale Reconstruction of the *Saccharomyces cerevisiae*. **Metabolic Network.**, p. 244 - 253, 2003.
- FORTINA, M. G., RICCI, G., MORA, D., PARINI, C., MANACHINI, P. L. Specific identification of *Lactobacillus helveticus* by PCR with *pepC*, *pepN* and *htrA* targeted primers. **FEMS Microbiol. Lett**, v.198, p.85-89, 2001.
- FOX, K. F., WUNSCH, D. S., FOX, D. A., STEWART, G. C. Complementarity of GC-MS and LC-MS analyses for determination of carbohydrate profiles of vegetative cells and spores of bacilli. **J. Microbiol. Methods**, v. 33, p.1-11, 1998.
- GARCÍA, P., ALLENDE, F., LEGARRAGA, P., HUILCAMAN, M., SOLARI, S. Identificación bacteriana basada en el espectro de masas de proteínas: Una nueva mirada a la microbiología del siglo XXI. **Revista chilena de infectología**, ed. 29, n. 3, pg. 263-272, 2012.

- GOMES, F. S. *Antagonismo entre leveduras e bactérias lácticas na fermentação etanólica*. Dissertação de Mestrado. 171f. Universidade de São Paulo, Escola Superior de Agricultura “Luiz de Queiroz”, Piracicaba, 2009.
- GOUY, M., GUINDON, S., GASCUEL, O. SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. **Mol Biol Evol**, v. 27, p. 221–224, 2010.
- GUARRO, J., GENÉ, J., STCHIGEL, A. M. Developments in Fungal Taxonomy. **Clinical microbiology reviews**, Washington , v.12, p. 454-500, 1999.
- HIMMELREICH, U., SOMORJAI, R. L., DOLENKO, B., DANIEL, H. M., SORRELL, T. C. A rapid screening test to distinguish between *Candida albicans* and *Candida dubliniensis* using NMR spectroscopy. **FEMS Microbiol Lett**, v. 251, p. 327–332, 2005.
- HJORT, K., LEMBKE, A., SPEKSNIJDER, A., SMALLA, K., JANSSON, J.K. Community structure of actively growing bacterial populations in plant pathogen suppressive soil. **Microbial Ecology**, New York, v. 53, p. 399-413, 2007.
- HOWELL, K. S., KLEIN, M., SWIEGERS, J. H., HAYASAKA, Y., ELSEY, G. M., FLEET, G. H., HØJ, P. B., PRETORIUS, I. S., DE BARROS LOPES, M. A. Genetic determinants of volatile-thiol release by *Saccharomyces cerevisiae* during wine fermentation. **Applied and environmental microbiology**, v.71.9, p. 5420-5426, 2005.
- KARAS, M., BACHMANN, D., HILLENKAMP, F. Influence of the Wavelength in high irradiance ultraviolet laser desorption mass spectrometry of organic molecules. **Anal. Chem.**, p.2935 -2939, 1985.
- KLIJN, N., WEERKAMP, A. H., DEVOS, W. M. Identification of mesophilic lactic acid bacteria by using polymerase chain reaction-amplified variable regions of 16S rRNA and specific DNA probes. **Applied and Environmental Microbiology** v.57, p. 3390–3393, 1991.
- KURTZMAN, C. P., ROBBETT, C. J. Identification and phylogeny of ascomycetous yeasts from analysis of nuclear large subunit (28S) ribosomal DNA partial sequences. **Antonie van Leeuwenhoek** , v.73, p.331–371, 1998.
- LAUDERBACH, A., UBESCK, J. C., BERH, J., VOGEL, R. F. MALDI-TOF MS typing enables the classification of brewing yeasts of the genus *Saccharomyces* to major beer styles. **PLoS One.**, v.12, n.8, 2017.
- LUCENA, B. T. L., DOS SANTOS, B. M., MOREIRA, J. L., MOREIRA, A. P. B., NUNES, A. C., AZEVEDO, V. Diversity of lactic acid bacteria of the bioethanol process. **BMC Microbiol.**,v.10, p.298, 2010.
- LONGO, M. A., NOVELLA, I. S., GARCIA, L. A., DIAZ, M. Comparison of *Bacillus subtilis* and *Serratia marcescens* as protease producers under different operating conditions. **J. Biosci. Bioeng.**, v. 88, p.35-40, 1999.
- LUDWIG KM, OLIVA-NETO P, ANGELIS DE, D F: Quantification of *Saccharomyces cerevisiae* flocculation by contaminant bacteria from alcoholic fermentation. **Ciênc Tecnol Aliment**, 21:63-68, 2001.
- MADOROBA, E., STEENKAMP, E. T., THERON, J., SCHEIRLINCK, I., CLOETE, T. E., HUYS, G. Diversity and dynamics of bacterial populations during spontaneous sorghum fermentations used to produce ting, a South African food. **Systematic and Applied Microbiology**, v. 34, n.3, p. 227-234, 2011.
- MAGNANI1, G. S., DIDONET, C. M., CRUZ, L. M., PICHETH, C. F., PEDROSA, F. O., SOUZA, E. M. Diversity of endophytic bacteria in Brazilian sugarcane. **Genet. Mol. Res.** 9 (1): 250-258, 2010.
- MALDI BIOTYPER CA SYSTEM. Clinical application for identification of microorganisms. Disponível em <[https://www.bruker.com/fileadmin/user\\_upload/8-PDF-Docs/Separations\\_MassSpectrometry/Literature/Brochures/MALDI-biotyper-CA-Brochure\\_08-2014\\_ebook.pdf](https://www.bruker.com/fileadmin/user_upload/8-PDF-Docs/Separations_MassSpectrometry/Literature/Brochures/MALDI-biotyper-CA-Brochure_08-2014_ebook.pdf)> Acesso em 12 de fevereiro de 2019.

- de MAN, J. C., ROGOSA, M., SHARPE, M. E. A medium for the cultivation of lactobacilli. **J. Bacteriol.** 23:130, 1960.
- MEDEIROS, E. M., de ALMEIDA, V. L., THOMAZINI, R. M. E. Implantação de um sistema de gestão: aplicação em uma usina sucroenergética. XIX Congresso Brasileiro de Custos – Bento Gonçalves, Brasil, 12 a 14 de novembro de 2012.
- MENDES, R., PIZZIRANI-KLEINER, A. A., ARAÚJO, W. L., RAAIJMAKERS, J. M. Diversity of cultivated endophytic bacteria from sugarcane: genetic and biochemical characterization of Burkholderia cepacia complex isolates. **Appl. Environ. Microbiol.** p. 7259- 7267, 2007.
- MINISTÉRIO DA AGRICULTURA, PECUÁRIA E ABASTECIMENTO (MAPA). Produção brasileira de cana-de-açúcar, açúcar e etanol. Setor Sucroalcooleiro, 2018. Disponível em: <http://www.agricultura.gov.br/assuntos/sustentabilidade/agroenergia/arquivos-producao/001PRODUOBRASILEIRADECANADEACARACAREETANOL.pdf> Acesso em: janeiro de 2018.
- MOOTHOO-PADAYACHIE, A., KANDAPPA, H. R., KRISHNA, S. B. N., MAIER, T., GOVENDER, P. Biotyping *Saccharomyces cerevisiae* strains using matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS). **European Food Research and Technology**, v.236, n.2, p.351-364, 2013.
- MORS, W. B., RIZZINI, C. T., PEREIRA, N. A., DEFILIPPS R. A. Medicinal Plants of Brazil **Reference Publications, Inc.: Algonac, MI**, p. 501, 2000.
- MORI, K., YAMAZAKI, K., ISHIYAMA, T., KATSUMATA, M., KOBAYASHI, K., KAWAI, Y., INONE, N. AND SHINANO, H. Comparative sequence analysis of the genes coding for 16S rRNA of Lactobacillus casei related taxa. **International Journal of Systematic Bacteriology**, v.47, p.54–57, 1997.
- MOUSSAQUI, W., JAULHAC, B., HOFFMANN, A. M., LUDES, B., KOSTRZEWA, M., RIEGEL, P., PREVOST, G. Matrix-assisted laser desorption ionization time-of-flight mass spectrometry identifies 90% of bacteria directly from blood cultures vials. **Clinical Microbiology and Infection**, n.16, p. 1631 – 1638, 2010.
- NARENDRANATH, N. V., HYNES, S. H., THOMAS, K. C., INGLEDEW, W. M. Effects of Lactobacilli on Yeast-Catalyzed Ethanol Fermentations. **Applied and Environmental Microbiology**, v.63, n.11, p.4158-4163, 1997.
- NASER, S. M., THOMPSON, F. L., HOSTE, B., GEVERS, D., DAWYNDT, P., YANCANNEYT, M. SWINGS, J. Application of multilocus sequence analysis (MLSA) for rapid identification of Enterococcus species based on rpoA and pheS genes. **Microbiology**. Washington, v.151, pt. 7, p. 2141- 2150, 2005.
- NASER, S. M., DAWYNDT, P., HOSTE, B., GEVERS, D., VANDEMEULEBROECKE, K., CLEENWERCK, I., YANCANNEYT, M., SWINGS, J. Identification of *lactobacilli* by *pheS* and *rpoA* gene sequence analyses. **Int. J. Syst. Evol. Microbiol.** Reading UK, v. 57, pt. 12, p. 2777 – 2789, 2007.
- NAVES, R. F., FERNANDES, F. S., PINTO, O. G., NAVES, P. L. F. Contaminação microbiana nas etapas de processamento e sua influência no rendimento fermentativo em usina alcooleira. **Enciclopédia Biosfera**, v. 6, n. 11, p. 1-16, 2010.
- NEELAKANTAM, V., NARENDRANATH, N. V., POWER, R. Relationship between pH and medium dissolved solids in terms of growth and metabolism of Lactobacilli and *Saccharomyces cerevisiae* during ethanol production. **Appl Environ Microbiol**, v.71, p.2239-2243, 2005.
- NEVOIGT, E. Progress in metabolic engineering of *Saccharomyces cerevisiae*. **Microbiol Mol Biol Rev**, v.72, p.379–412, 2008.

- NGUYEN, D. T. L., VAN HOORDE, K., CNOCKAERT, M., de BRANDT, E., de BRUYNE, K., LE, B. T., VANDAMME, P. A culture-dependent and independent approach for the identification of lactic acid bacteria associated with the production of nem chua, a Vietnamese fermented meat product. **Food Research International**, v.50, p.232-240, 2013.
- NOBRE, T. P., HORII, J., ALCARDE, A.R. Cellular viability of *Saccharomyces cerevisiae* cultivated in association with contaminant bacteria of alcoholic fermentation. **Ciência Tecnol Aliment**, 27, p. 20-25, 2007.
- NOLASCO- JÚNIOR, J. 2010. *Eficiência de processo térmico para mostos a base de caldo de cana e melão na produção de bioetanol*. 221f. Tese de doutorado. Universidade Estadual de Campinas, Campinas, 2010.
- NOVA CANA.com. Produção de etanol: Tecnologia praticada no Brasil – 2018- <https://www.novacana.com/etanol/producao-tecnologia-praticada-no-brasil> - acesso em: 13 de dezembro de 2018.
- OKI, K., RAI, A. K., SATO, S., WATANABE, K. TAMANG, J. P. Lactic acid bacteria isolated from ethnic preserved meat products of the Western Himalayas. **Food Microbiology**, v.28, p.1308-1315, 2011.
- OLIVEIRA, K. Z. *Construção de árvores filogenéticas baseadas em genomas completos*. Dissertação de Mestrado. 101f. Universidade Estadual de Campinas, Campinas, 2010.
- OVIANO, M., SPARBIER, K., BARBA, M. J., KOSTRZEWA, M., BOU, G. Universal protocol for the rapid automated detection of carbapenem-resistant Gram-negative bacilli directly from blood cultures by matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry (MALDI-TOF/MS). **International Journal of Antimicrobial Agents**, La Coruña, v.48, p.655-660, 2016.
- PACE, N. R. A molecular view of microbial diversity and the biosphere. **Science**, Washington, v. 276, p. 734-740, 1997.
- PASSOTH, V., BLOMQUIST, J., SCHNÜRER, J. *Dekkera bruxellensis* and *Lactobacillus vini* form a stable ethanol-producing consortium in a commercial alcohol production process. **Appl Environ Microbiol.**, 73:4354-4356, 2007.
- PETERSOHN, A., BRIGULLA, M., HAAS, S., HOHEISEL, J. D., VÖLKER, U., HECKER, M. Global analysis of the general stress response of *Bacillus subtilis*. **J. Bacteriol.**, v.183, p.5617-5631, 2001.
- PLONGLA, R., PANAGEA, T., PINCUS, D. H., JONES, M. C., GILLIGAN, P. H. Identification of *Burkholderia* and uncommon glucose-nonfermenting Gram-negative bacilli isolated from patients with cystic fibrosis by use of matrix-assisted laser desorption ionization–time of flight mass spectrometry (MALDI-TOF MS). **J Clin Microbiol**, v. 54, p.3071–3072, 2016.
- PODLESNY, M., JAROCKI, P., KOMON, E., GLIBOWSKA, A., TARGONSKI, Z. LC-MS/MS Analysis of Surface Layer Proteins as a Useful Method for the Identification of *Lactobacilli* from the *Lactobacillus acidophilus* Group. **J. Microbiol. Biotechnol.** Lublin. p. 421-429, 2011.
- RANJARD, L., POLY, F., NAZARET, S. Monitoring complex bacterial communities using culture-independent molecular techniques: application to soil environment. **Research in Microbiology**, Paris, v. 151, p. 167-177, 2000.
- RODAS, A. M., CHENOLL, E., MACIAN, M. C., FERRER, S., PARDO, I., AND AZNAR, R. **Int. J. Syst. Evol. Microbiol.**, v.56, p. 513–517, 2006.
- SAITOU, N., NEI, M. The neighbour-joining method: a new method for reconstructing phylogenetic trees. **Mol. Biol. Evol.**, v.4, 406–425, 1987.

- SANDHU, G. S., KLINE, B. C., STOCKMAN, L., ROBERTS, G. D. Molecular Probes for Diagnosis of Fungal Infections. **J. Clin.Microbiol.**, v. 33, n.11, p. 2913 -2919, 1995.
- SANDRIN, T. R., GOLDSTEIN, J. E., SCHUMAKER, S. MALDI TOF MS profiling of bacteria at the strain level: A review. *Mass Spectrom Rev.* v.32, (3), p.188-217, 2013.
- SAUER, S., KLIEM, M. Mass spectrometry tools for the classification and identification of bacteria. **Nat. Ver. Microbiol.**, v.8, n.1, p. 74-82, 2010.
- SCHOCH, C. L., SEIFERT, K. A., HUHNDORF, S., ROBERT, V., SPOUGE, J. L., LEVESQUE, C. A., CHEN, W., FUNGAL BARCODING CONSORTIUM. Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for *Fungi*. **Proceedings of the National Academy of Sciences**, v.109, n.16, p.6241-6246, 2012.
- SCORZETTI, G., FELL, J. W., FONSECA, A., STATZELL-TALLMAN, A. Systematics of basidiomycetous yeasts: A comparison of large subunit D1/D2 and internal transcribed spacer rDNA regions. **FEMS Yeast Res**, v.2, p.495–517, 2002.
- SENG, P., DRANCOURT, M., GOURIET, F., LA SCOLA, B., FOURNIER, P. E., ROLAIN, J. M., RAOULT, D. Ongoing revolution in bacteriology: routine identification of bacteria by matrix-assisted laser desorption ionization time-of-flight mass spectrometry. **Clin. Infect. Dis.**, v.49, n.4, p.543-551, 2009.
- SEGAWA, S., SAWAI, S., MURATA, S., NISHIMURA, M., BEPPU, M., SOGAWA, K., WATANABE, M., SATOH, M., MATSUTANI, T., KOBAYASHI, M., IWADATE, Y., KUWABARA, S., SAEKI, N., NOMURA, F. Direct application of MALDI-TOF mass spectrometry to cerebrospinal fluid for rapid pathogen identification in a patient with bacterial meningitis. **Clinica Chimica Acta**, v.435, p. 59-61, 2014.
- SCHIOZER, A. L., CABRAL, E. C., de GODOY, L. A. F., CHAVES, F. C. M., POPPI, R. J., RIVEROS, J. M., EBERLIN, M. N., BARATA, L. E. S. Electrospray Ionization Mass Spectrometry Fingerprinting of Extracts of the Leaves of *Arrabidaea chica*. **J. Braz. Chem. Soc.**, v.23, n.3, p. 1592–1602, 2012.
- SHANDU G. S., KLINE, B. C., STOCKMAN, L., ROBERTS, G. D. Molecular probes for diagnosis of fungal infections **J. Clin.Microbiol.** v.33, n.11, p.2913, 1995.
- SILVA, D. M., CARVALHO, K. A., MOURA, A. A. O. 2017 Contaminação bacteriana na fermentação e sua influência na eficiência industrial de uma usina de cana-de-açúcar. Trabalho de conclusão de curso, Universidade de Rio Verde, 2017.
- SILVA-FILHO, E. A., MELO, H. F., ANTUNES, D. F., SANTOS, S. K. B., RESENDE, A. M., SIMÕES, D. A., MORAIS, M. A. Jr: Isolation by genetic and physiological characteristics of a fuel-ethanol fermentative *Saccharomyces cerevisiae* strain with potential for genetic manipulation. **J Ind Microbiol Biotechnol.**, v. 32, p.481-486, 2005.
- SINGH, S., GOSWAMI, P., SINGH, R.,HELLER, K. J. Application of molecular identification tools for *Lactobacillus*, with a focus on discrimination between closely related species: a review. **LWT – Food Science and Technology**, v.42, p. 448–457, 2009.
- SINGHAL, N., KUMAR, MANISH., KANAUIA, P. K., VIRDI, J. S. MALDI-TOF mass spectrometry: an emerging technology for microbial identification and diagnosis. **Front. Microbiol.**, v.6, 2015.
- SOBRINHO, V. S., SILVA, V. C. F., CEREDA, M. P. Fermentation of sugar cane juice (*Saccharum officinarum*) cultivar RB 7515 by wild yeasts resistant to UVC. **Journal of Biotechnology and Biodiversity**. v. 2, n.1, p. 13-21, 2011.

- de SOUSA, J. L. U., MONTEIRO, R. A. B. Fatores interferentes na fermentação alcoólica para a produção de etanol. **FAZU em Revista**. v. 2, p. 1-8, 2011.
- de SOUZA, R. B., DOS SANTOS, B. M., DE SOUZA, R. F. R., DA SILVA, P. K., LUCENA, B. T., DE MORAIS, M. A. JR. The consequences of *Lactobacillus vini* and *Dekkera bruxellensis* as contaminants of the sugarcane-based ethanol fermentation. **J. Ind. Microbiol Biotechnol.**, v.39, p.1645-1650, 2012.
- STIRLING, D. DNA extraction from fungi, yeast, and bacteria. In **PCR protocols**, p. 53-54, 2003.
- SUE, T., OBOLONKIN, V., GRIFFITHS, H., VILLAS-BÔAS, S. G. Monitoring microbial contamination in microalgae fermentation processes using metabolic footprint analysis – an exometabolomics approach. Washington: **Applied and Environmental Microbiology**, p. 7605- 7610, 2011.
- SUGARCANE. Org Brazil's Diverse Energy Matrix. Site disponível em: < <http://sugarcane.org/the-brazilian-experience/brazils-diverse-energy-matrix>> Acesso em 12 dezembro de 2017.
- TIUKOVA, I., EBERHARD, T., PASSOTH, V. Interaction of *Lactobacillus vini* with the ethanol-producing yeasts *Dekkera bruxellensis* and *Saccharomyces cerevisiae*. **Biotechnology and Applied Biochemistry**, p.40-44, 2014.
- VASQUEZ, A., MOLIN, G., PETERSON, B., ANTONSSON, M., AHRNÉ, S., DNA- based classification and sequence heterogeneities in the 16S Rdna genes of *Lactobacillus casei/ paracasei* and related species. **Syst. Appl. Microbiol.** New York, v.28, no 5, p. 430-441, 2005.
- VILLAS-BÔAS, S. G., MAS, S., AKESSON, M., SMEDSGAARD, J., NIELSEN, J. Mass spectrometry in metabolome analysis. **Mass Spectrometry Reviews**, v.24, p. 613–646, 2005.
- de VUYST, L., VANCANNEYT, M. Biodiversity and identification of sourdough lactic acid bacteria. **Food Microbiol.** Brussels.v.24, p. 120-127,2007.
- TIUKOVA, I., EBERHARD, T., PASSOTH, V. Interaction of *Lactobacillus vini* with the ethanol-producing yeast *Dekkera bruxellensis* and *Saccharomyces cerevisiae*. **Biotechnol. Appl. Biochem.**, v.61, p.40-44, 2014.
- TRINGE, S., HUGENHOLTZ, P. A renaissance for the pioneering 16S rRNA gene. **Current Opinion in Microbiology**, Oxford, v. 11, p. 442-446, 2008.
- YOKOTA F, OLIVA NETO P: Características da flocculação de leveduras por *Lactobacillus fermentum*. **Rev Microbiol**, v.22, p.12-16, 1991.
- WAHL, K. L., WUNSCH, S. C., JARMAN, K. H., VALENTINE, N. B., PETERSEN, C. E., KINGSLEY, M. T., ZARTOLAS, K. A., SAENZ, A. J. Analysis of microbial mixtures by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. **Anal. Chem.**, v.74, p.6191-6199, 2002.
- WARD, L. J., TIMMINS, M. J. Differentiation of *Lactobacillus casei*, *Lactobacillus paracasei* and *Lactobacillus rhamnosus* by polymerase chain reaction. **Lett Appl Microbiol**, v.29, p.90-92, 1999.
- WEST, E. R., COTHER, E. J., STEEL, C.C., ASH, G. J. The characterization and diversity of bacterial endophytes of grapevine. **Canadian Journal of Microbiology**, Ottawa, v. 56, p. 209-216, 2010.
- Wilkins, C. L., Lay, J. O. Identification of microorganisms by mass spectrometry, In: **Winefordner JD (ed)**, vol 169. Wiley, New Jersey, 2006.
- WU, Z., WANG, X. R., BLOMQUIST, G. Evaluation of PCR primers and PCR conditions for specific detection of common airborne fungi, v.4, n.3, p. 377-382, 2002.
- ZEIGLER, D. R. Gene sequences useful for predicting relatedness of whole genomes in bacteria. **International Journal of Systematic and Evolutionary Microbiology**, v. 53, p. 1893-1900, 2003.

ZHANG, L., VRANCKX, K., JANSSENS, K., SANDRIN, T.R. Use of MALDI-TOF Mass Spectrometry and a Custom Database to Characterize Bacteria Indigenous to a Unique Cave Environment (Kartchner Caverns, AZ, USA). **J. Vis. Exp.**, v.95, p.1-12, 2015.